DOCUMENT RESUME

ED 297 006 TM 011 952

AUTHOR Ingels, Steven J.; And Others
TITLE Symposium on the National Education Longitudinal Study of 1988 (NELS:88) and the NELS:88 Field Test (New Orleans, Louisiana, April 5-9, 1988).
PUB DATE Mar 88
NOTE 109p.; For the two papers by R. P. Russo, see ED 293 893 and ED 295 985.
PUB TYPE Collected Works - Conference Proceedings (021) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Field Tests; Grade 8; Junior High Schools; *Junior High School Students; *Longitudinal Studies; *National Surveys; *Outcomes of Education; Parent Student Relationship; Research Methodology; School Surveys; Student Attitudes; Teacher Attitudes; Test Construction; Test Validity; Trend Analysis
IDENTIFIERS *National Education Longitudinal Study 1988; Parent Surveys; Student Surveys; Teacher Surveys

ABSTRACT

The National Education Longitudinal Study of 1988 (NELS:88) is a major new panel study of educational outcomes sponsored by the Center for Education Statistics of the United States Department of Education. The NELS:88 is designed to provide trend data about critical transitions experienced by young people as they develop, attend school, and embark on their careers. Initially focusing on eighth graders, the cohort will be reassessed at two-year intervals. Six papers delivered during this symposium include: (1) "The National Education Longitudinal Study of 1988: NELS:88, Its Design and Objectives" (S. J. Ingels); (2) "The National Education Longitudinal Study of 1988 School Survey" (R. P. Russo); (3) "The National Education Longitudinal Study of 1988 Teacher Survey" (R. P. Russo); (4) "The Design and Validation of the NELS Test Battery" (D. A. Rock and J. M. Pollack); (5) "Methodological Experiments of the NELS:88 Field Test Student Survey" (S. R. Lucas); and (6) "Methodological Experiments of the NELS:88 Field Test Parent Survey" (B. B. Simon). (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED297006

SYMPOSIUM ON THE NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988 (NELS:88) and THE NELS:88 FIELD TEST

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
S. INGELS
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

American Educational Research Association Meetings,
New Orleans, Louisiana, April 1988

contributors --

Steven J. Ingels, NORC
Donald Rock and **Judith Pollack**,
Educational Testing Service
Rocco Russo, Westat
Samuel Lucas, NORC
Betty Simon, NORC

chair --

Anne Hafner, Longitudinal Studies Branch,
Center for Education Statistics,
U.S. Department of Education

discussant --

John Stiglmeier, New York State Department of Education

TM011952

SYMPOSIUM ON THE NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988 (NELS:88) and THE NELS:88 FIELD TEST

I. INTRODUCTION: NELS:88, ITS DESIGN AND OBJECTIVES

Steven J. Ingels, NORC

II. THE NELS:88 SCHOOL SURVEY

Rocco P. Russo, Westat, Inc.

III. THE NELS:88 TEACHER SURVEY

Rocco P. Russo, Westat, Inc.

IV. THE DESIGN AND VALIDATION OF THE NELS:88 TEST BATTERY

Donald A. Rock and Judith Pollack, ETS

V. STUDENT SURVEY: METHODOLOGICAL EXPERIMENTS IN THE NELS:88 FIELD TEST

Samuel R. Lucas, NORC

VI. PARENT SURVEY: METHODOLOGICAL EXPERIMENTS IN THE NELS:88 FIELD TEST

Betty B. Simon, NORC

THE NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988:

NELS:88, ITS DESIGN AND OBJECTIVES

Steven J. Ingels

NORC
A Social Science Research Center
University of Chicago
1155 E. 60th Street
Chicago, Illinois 60637

Introduction to the Symposium on NELS:88 and the NELS:88 Field Test, American Educational Research Association Annual Meetings, New Orleans, Louisiana, April 1988.

Steven J.Ingels, NORC

NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988

Symposium on NELS:88 and its National Field Test

AERA, New Orleans, April 8, 1988

Division D; SIC/Longitudinal Research

INTRODUCTION: NELS:88, ITS DESIGN AND OBJECTIVES

Overview of the Study Design

The National Education Longitudinal Study of 1988 is a major new panel study of educational outcomes sponsored by the Center for Education Statistics of the U.S. Department of Education. It is designed to provide trend data about critical transitions experienced by young people as they develop, attend school, and embark on their careers. By initially focusing in 1988 on eighth graders and their schools, teachers, and parents, then by following up that cohort at two-year intervals, the NELS:88 research will address such crucial research questions as:

-- What are the determinants of persistence in and completion of high school? What types of programs are being used to prevent dropping out, and which succeed? What are the differences between earliest and later dropouts, especially in regard to educational experiences?

-- What are the transition patterns from eighth grade to secondary school? How much movement across "sectors" (public to private; private to public) occurs at this point? What is the impact of moving from more socially homogeneous neighborhood schools to more heterogeneous consolidated schools?

-- What are the dynamics of tracking? How soon and by what mechanisms are students grouped into educational program tracks? What are the roles of students, teachers, administrators and parents in that process?

-- What are the features of effective schools? Do the features differ for particular groups of children? What is the impact upon school effectiveness of the size and characteristics of the faculty and staff? How do such factors as staff professionalism, racial-ethnic composition, and degree of teacher specialization, singly and jointly affect school effectiveness?

For the base year survey (1988), a nationally representative sample of 1,000 schools (800 public schools; and 200 private schools, including parochial institutions) has been drawn. Within this school sample, 26,200

eighth grade students have been selected at random. Base-year data collection is occurring at this time. A base-year data file and user's manual will be available in the summer of 1989. Findings from the study will be published periodically beginning in the spring of 1989.

The Four Component Surveys: Students, Parents, Teachers, Administrators

The overall design for the base-year (1988) study includes four component surveys--students, parents, school administrators and teachers. A student questionnaire will gather information about basic background variables and a range of additional topics including school work, aspirations and social relations. The students will also complete a cognitive test battery consisting of a curriculum-sensitive test that will use item overlap methods to measure cognitive growth between eighth and twelfth grades in four achievement areas: reading and vocabulary; mathematics, science; and social studies. A school administrator questionnaire will be completed by school principals and will gather descriptive information about the school's teaching staff, student body, school polices and offerings, and will ascertain which high schools the school's students are likely to attend. Selected teachers (in two of the four subject achievement areas) of each sampled child will complete a teacher questionnaire which will collect data about school and teacher characteristics, evaluations of the selected students, and classroom teaching practices. One parent of each student will be asked to respond to a parent survey, intended to gauge parental aspirations for their children, family willingness to commit resources to children's education, the home educational support system, and other family characteristics relevant to achievement.

The four questionnaires and the cognitive tests each form the topic for a separate paper to be presented in this symposium. Below, we try to sketch the sample design and some of the policy analysis goals of the study as a point of entry into the discussion of the four component surveys of students, teachers, parents and principals, and of the research instruments developed for each respondent population: the four NELS:88 questionnaires and the test battery.

Base Year Sample Design

The NELS:88 Base Year sample is representative of eligible eighth grade schools and students in the 50 states plus the District of Columbia. A sample of public and private schools was drawn with probabilities proportionate to eighth grade enrollment in June of 1987.

Specifically, students are being sampled through a two-stage process, modelled after that used for the two prior CES longitudinal surveys, NLS-72 and High School and Beyond. The first stage involved stratified sampling of some 1,000 public and private schools from a universe of approximately 40,000 schools containing eighth grade students. Stratification is by administrative control, with an oversample of Catholic and other private schools; and by geography, permitting comparisons among the nine Census

divisions. The sample will provide state-representative and comparative data for the largest states (California, Florida, Illinois, Michigan, New York, Ohio, Pennsylvania, and Texas) and for smaller states that have chosen to augment their samples.

The second stage includes random selection of about 26 (24 core, and, on average, 2.2 oversampled Hispanic and Asian supplemental) students per school. Once students are selected, the NELS:88 sample design includes one parent for each selected student, and two teachers in designated subject areas. Follow-up surveys are planned for 1990 and 1992, when most of the initial cohort will be in the tenth (1990) and twelfth (1992) grades, and 1994. Students who drop out of school during this period will also be surveyed.

One parent of each child will be included in the study. Parent data will be used primarily in the analysis of student behaviors and outcomes, and only secondarily as a dataset by itself. This purpose does not require an independent probability sample of parents, but rather the collection of data from the most knowledgeable parent or guardian of each student.

All full- and part-time teachers who are teaching classes in mathematics, science, English/language arts, and social studies to eighth graders in the spring of 1988 are included in the NELS:88 universe of eighth grade teachers. The actual sample is restricted to teachers who are providing instruction in the listed subjects to the selected sample of eighth grade students within the sampled schools. Two teachers are being sampled for each selected student. The administrative head of each school is also included in the sample. Again, since the student is the unit of analysis, this sample is not representative of eighth grade teachers, but of selected teachers of a nationally representative sample of eighth graders.

The NELS:88 design consists of a core sample, plus a supplementary sample of students (and parents and teachers) sponsored by the Office of Bilingual Education and Minority Language Affairs (OBEMLA). The OBEMLA supplement includes additional students of Hispanic and Asian descent beyond the numbers selected for the core study. This oversample of Hispanic and Asian students is required to obtain sufficient numbers for these groups and their principal subgroups to support separate policy analyses.

Several states have chosen to augment their samples, adding schools in order to provide a state representative sample (Vermont, Washington, and Connecticut) or to extend the range and statistical precision of the comparisons and analyses that the NELS:88 sample can sustain (New York). In addition, Gallaudet University has sponsored a special oversample of hearing impaired students who are mainstreamed in eighth grade English or mathematics. These sample augmentations will add approximately 4,000 students to the study.

Analytical Utility of the NELS:88 Program

The basic design for NELS:88 is longitudinal in nature. This longitudinal study is part of a larger set of longitudinal research activities which are mandated by Congress and carried out by the Center for Education Statistics (CES). Indeed, it is the third in a series of C.E.S. longitudinal studies, its predecessors being the National Longitudinal Study of the High School Class of 1972 (NLS-72) and High School and Beyond (HS&B). The study, by beginning with a cross-section of 1988 eighth graders, following a substantial subsample of these students in 1990 and thereafter, and by freshening the 1990 sample, will provide a point of comparison with the high school class of 1982, studied by HS&B, and the high school class of 1972, studied by NLS-72.

The overall scope and longitudinal design of the study offer at least four major benefits:

A. It provides the basis for within-cohort comparison by following the same individuals over time.

Many, if not most, of the questions policymakers wish to answer involve some notion of change over time: Under what conditions do students shift their educational choices? Are there any peculiar factors at work (e.g., tougher educational standards) that explain changes in the educational performance of students at the aggregate level? What is the role of families in the shaping of attitude and school behavior?

Certainly cross-sectional analysis may approximate the study of the process of change by using a number of devices, but the risk involved is large and the possibility of examining causal relationships is nil. For example, a study may ask respondents who have failed to complete their elementary education when they started to think about or plan on dropping out, but the danger here is that the farther back they are forced to reach into their memories, the less accurate they are likely to be. A single study may indicate that younger adults are more likely to drop out than older ones, but this approach does not allow the examination of the causal link between the intention to drop out and actual dropping out.

B. It provides a representative cross-section of eighth graders in the United States, thereby shedding light on the factors at the demographic and environmental levels that affect educational outcomes such as school performance, individual aspirations, and career choices.

C. It concentrates on early adolescence, a period in which major changes in attitudes and behaviors take place. Thus it provides a basis for understanding such current areas of concern as the impact of curriculum content on educational achievement and occupational choices.

D. As part of the Longitudinal Studies Program of the Center for Education Statistics and as a successor to the NLS-72 and HS&B surveys, NELS:88 offers the opportunity for the analysis of trends in areas such as academic performance.

Taken together, the surveys in this program not only can inform policymakers about differences in performance among those who are 13, 15, or 17 years of age, but can also provide them with data for the assessment or evaluation of programs geared to improve the quality of education offered by the schools.

Policy Relevance of the NELS:88 Data: Areas of Inquiry

NELS:88's stated objectives are far more comprehensive than those for any education longitudinal study to date. Its major features--the planned integration of school, student, parent and teacher studies; the initial concentration on eighth grade student cohorts with planned follow-up at two year intervals; the inclusion of supplementary components to support analyses of geographically or demographically distinct subgroups; and the design linkages to previous longitudinal studies and other current studies--stamp it with an outstandingly ambitious research design.

The major objectives of NELS:88 are to serve the development and evaluation of educational policy at all governmental levels; and to inform decision-makers, educational practitioners, and parents about the changes in the operation of the educational system across time, and the effects of various elements of the system on the lives of the individuals who pass through it. The most important policy issues to be addressed by NELS:88 concern the patterns of transition from eighth grades in elementary, middle, or junior high schools to tenth grades in secondary schools.

These patterns of transition cover individual as well as institutional characteristics. At the individual level the study will look into educational attainment and personal development. In response to policy and scientific issues, data will also be provided on the demographic and background correlates of educational outcomes. At the institutional level, NELS:88 will focus on school "effectiveness" issues such as tracking, promotion, retention and curriculum content and structure.

Finally, this study will focus attention on basic policy issues related to parents' role in the educational success of their children. The extent to which schools are expected to provide special services to selected groups of students to compensate for limitations in the home; the effect of parents' attitudes on educational choices; the correlates of active parental involvement in the school--these are some of the questions to be addressed by NELS:88.

Thus the base year and follow-up studies will explore a number of areas that define the basic outcome variables of the study--those related to educational achievement, occupational expectations and achievement, and personal and social development. Information will be gathered as well on numerous independent variables, such as standard demographics, and variables measuring educational support, parent's socioeconomic status, family composition, language use, and home environment. The core of intervening variables encompasses school experiences such as exposure to given curriculum content and structure, assessment and evaluation systems, social relations, school behavior, and participation in extracurricular activities.

Each stage in the collection of data will gather information on relevant experiences in order to build the necessary basis for a longitudinal analysis. Emphasis on specific areas of inquiry will change, however, in response to the particular stages of development that respondents will be in at a particular point in time. For example, the 1988 survey cautiously approaches the question of occupational expectations, anticipating these to be less than fully crystallized in a group whose modal age is 13 years. It is likely that tenth graders will have a clearer understanding of the intricacies of the job market and the matching of their abilities with the opportunities available than will their younger counterparts in eighth grade. Thus, expanded coverage of this research area is planned for the 1990 survey.

Although the most powerful longitudinal analyses supported by NELS:88 will not be possible until several years after its inception, important cross-sectional analyses and descriptive data will be available from each survey wave. The discussion below thus will include research questions about static conditions as well as questions about stability and change.

Educational and Cognitive Outcomes

1. Educational Outcomes

NELS:88 will provide data needed to investigate the determinants of school achievement and completion. Major questions include what types of programs are being used to prevent dropping out, and which succeed. What are the differences between earliest dropouts and those who leave later in terms of social and demographic differences, retention in grade prior to eighth grade, differential likelihood of returning later to complete high school, occupational outcomes, etc.? What are the differences among racial/ethnic groups in the decision to leave school before the tenth grade?

2. Cognitive Outcomes

An achievement test designed by the Educational Testing Service (ETS) will be the basis for carrying out longitudinal comparisons of growth in achievement at both the individual and aggregate level. The battery will cover four achievement areas-- mathematics, reading, science and social studies. It will retain sufficient comparability with tests given to the 1972 NLS-72 seniors and the 1980 HS&B sophomore and senior cohorts. That

is, in addition to measuring within student changes, the scale will be able to support cross-sectional contrasts.

Analysis will focus on schooling experiences that affect subsequent cognitive measures. Information gathered in this area will touch upon several policy-relevant issues: To what extent does the participation in programs for learning assistance in a particular subject affect scores in that area? How much exposure to these programs is needed before any evidence of a positive impact comes to light? With base year measures held constant, what is the differential cognitive development of students with similar characteristics who are in schools where the "mainstreaming" approach is prevalent, versus those who are in schools that group students according to their individual abilities?

Occupational Outcomes

Analysis here will focus on the evolution of occupational choices, the impact of exposure to vocational training on educational choice, and successful entry into the labor market.

Personal and Social Development

The set of variables on self-esteem, values, and educational aspirations will provide answers to questions such as to what extent does academic success have longterm effects on subsequent psychological well-being and social adjustments? To what extent do family and school characteristics shape self-confidence and aspirations?

Independent Variables

In order to provide the grounds for explanatory or causal analyses, the NELS:88 design includes the measurement of a number of antecedent factors known to be relevant to the explanation of the outcome variables of the study. Demographic characteristics and educational support factors constitute the two sets of independent variables to be measured.

1. Demographic Variables

A large number of studies have shown that race, ethnicity, and socio-economic status are important predictors of educational outcomes and personal and social development. No large scale study of educational achievement and performance can therefore ignore the contribution of these "given" factors to variations in its major dependent variables. No matter how strong the correlations between demographic factors and educational outcomes, these "given" factors explain only partially the variance in the dependent variables. Another crucial group of variables for both policy and scientific purposes relates to the educational support that families give their children in pursuit of educational goals.

2. Educational Support

The educational support system covers a wide range of factors such as parental involvement in school, parental educational aspirations for their children, financial resources applied to educational expenses, and size of family. Important questions are: What are the correlates of parents' active involvement in their children's schooling? Do children of families who have higher educational expectations perform better than those who do not? How do parental roles differ in the case of single-parent families or families with both parents working? How do family size and birth order affect a child's school performance? What is the confluence between the values of parents and those of their eighth grade children in regard to education and work?

Intervening Variables

It is of practical and theoretical interest to discover what it is about a particular environment that makes it effective in a particular way. NELS:88 intends to go beyond the identification of school environment effects and seek an explanation of such effects. Such explanation requires not only the measurement of characteristics of eighth, tenth and twelfth graders but also the measurement of characteristics of the school environment.

The NELS:88 design achieves this goal in two ways: First, it incorporates in the student instrument a series of measures on exposure to curriculum content, special programs and perceptions of school climate.

Second, it collects supplemental information from principals on school characteristics such as grading and testing structure, school programs, school-parent interactions and general student and staff characteristics. These data will constitute the basis for analyses of school and specific intervening programs effects at the aggregate level.

A third set of variables will look at the fit between individual and home and school environments. They include measures on school-related behaviors and leisure-time activities whose impacts on educational achievements, values, and aspirations have not yet been sufficiently investigated.

The 1987 Field Test

NELS:88 instruments and procedures were field-tested in 52 eighth grade schools between February and May of 1987. This pretest was conducted in five states -- New York, Illinois, Florida, Texas, and California -- chosen on the basis of their regional representativeness and demographic heterogeneity. (A full report of field test results is given in Ingels et al., 1987.)

The development of the NELS survey instruments and forms was subject to a number of constraints, in the form of competing goals and limitations of instrument length. Since NELS:88 was designed to collect data from multiple sources, and because respondent burden must be limited, a leading principle of instrument design was that items in each questionnaire should generally be limited to those for which the targeted population constitutes the best respondent or data source. Thus data on students' home and learning environments should be collected primarily from parents and school administrators and teachers. Data on individual students' experiences, perceptions, beliefs, expectations and attitudes should be collected primarily from students. Other kinds of data might best be gathered from school records. Of course, some content redundancy between questionnaires is desirable to permit analyses of item validities and significant perceptual differences. In the main, however, a guiding intuition for questionnaire development was that each questionnaire should have the highest degree of articulation with other instruments, but should exhibit minimal repetition of survey items more appropriate for other respondents.

Development of the student questionnaire was subject to other design constraints. Since the foremost purpose of NELS:88 is to inform educational policy, highest priority was put on collecting policy-relevant data on educational outcomes and processes.

In addition, since the primary research objectives of NELS:88 are longitudinal in nature, priority was placed on developing survey items expected to be most useful for predicting or explaining future outcomes as measured in later survey waves, rather than items seeking retrospective information to explain students' current status. Likewise, in respect of the longitudinal objectives of the test battery to measure students' cognitive growth between the eighth, tenth, and twelfth grades in the academic subjects of mathematics, science, reading/language arts, and social studies, priority was placed on developing a dimensionally stable battery of four tests, suitable for use with all three grade levels, starting with eighth graders in 1988.

Since NELS:88 is the third installment in a continuing program of nationally representative education longitudinal studies, the tests and questionnaires had also to contain items that support cross-cohort comparisons and trend analyses, using the prior CES longitudinal surveys of the high school classes of 1972, 1980 and 1982. Thus the NELS:88 instruments needed to cover many of the same topic areas as NLS-72 and HS&B --to the extent that these are relevant to a somewhat younger population-- and when possible, with the same items or item wordings--to the extent consistent with comprehension on the part of the younger and far more heterogeneous eighth grade population. At the same time as providing a point of continuity for comparisons with earlier cohorts, NELS:88 needed to reflect changes registered in the last two decades in policy issues and their conceptualization. In particular, as a study of school effects as well as individual level behavior and outcomes, NELS:88 instruments needed to collect data concerning the effects of recent school reform initiatives, a paradigm case of the kind of intervention the effects of which a longitudinal study is best suited to detect.

Pretests, notoriously, have but modest predictive utility, in terms of main survey response rates and other results. But they may have enormous analytical value, particularly when selectively focused on the unique questions that the configured elements of a particular study present. Several questions assumed particular salience in the NELS:88 field test. These, primarily, were (1) questions of adequate means for the measurement of change over time -- as in the case of the cognitive battery; (2) questions of the proper respondent population to which a question might be posed, as in the issue of the extent to which eighth graders, or their parents, should be relied upon for basic SES data; (3) and questions of the implications of interaction of survey methodology with respondent definition, as in the case of the parent survey, where prior decisions limited response to one parent of each selected child, but did not prespecify in terms of gender which parent should be surveyed, leaving this matter, rather, to a process of self-selection that ideally would reflect a familial consensus on which parent was most knowledgeable about the child's schooling and home educational support system.

The uses of the field test in determining the final version of the NELS:88 instruments and procedures are described in the five component-centered papers that succeed this introduction to the study as a whole. In these, Rocco Russo addresses content and data collection issues pertaining to the administrator (principal/headmaster) and teacher questionnaires; Don Rock and Judy Pollack describe the development and validation of the cognitive test battery; and Betty Simon and Sam Lucas report on methodological experiments in the field test student and parent surveys.

Rocco identifies the purposes and objectives of the school (principal) survey, describes survey administration and data preparation activities, and presents results of the field test. Rocco depicts the design of the teacher study -- describing how teachers will be sampled and teacher-student data linked -- and reports on data production activities and field test conclusions.

Don and Judy focus on the issue of measuring group and individual change over time, comparing the complementary but contrasting aims of cross-sectional assessment designs (such as NAEP) and the NELS panel design, and the different kinds of policy analysis these designs support. They describe the principles of item selection and validation for the four NELS:88 cognitive tests, using the eighth, tenth, and twelfth grade data collected in the field test.

Sam raises the issue of who is the better respondent -- parent or child -- for certain kinds of information, by examining item nonresponse on items common to the student and parent questionnaires. The word "experiment," I might add, must be used loosely in a school-based survey where a random sample of schools will cooperate under conditions that they individually choose and in environments that they largely control and that the experimenter must adapt to. The sort of controlled inquiry in which the experimenter is able to manipulate at will the relevant features, factors or variables assumed to be relevant conditions for the phenomenon under study can often only imperfectly be achieved in survey experiments in such

settings. This point, I think, will be underlined by Betty's analysis of the experiment built into the distribution of the parent questionnaire. We wanted to compare mail distribution with student take-home distribution of the parent survey instrument. However, certain schools selected themselves into additional variations on the distributive comparison: some would not release parent addresses, so questionnaires had to be mailed by the school. Other schools did not hold the parent questionnaire until Survey Day as instructed, but distributed them well in advance of the school's survey session -- with excellent results, by the way, but, given the small and self-selected number of schools, with results that could not be generalized from. A further interesting feature of the parent survey was that it had no predesignated parent respondent. Betty reports on who (which parent -- natural mother, or father; or step-parent; or other) actually answered the items in the parent questionnaire, and in so doing raises for consideration a basic design issue -- is one parent respondent an adequate representation of such constructs as "the parents," or "the home environment"?

Reference:

Ingels, S.J. et al., Field Test Report, National Education Longitudinal Study of 1988 (NELS:88). NORC report to Center for Education Statistics, Chicago, July 1987.

THE NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988 SCHOOL SURVEY*

Rocco P. Russo

Westat, Inc.
1650 Research Boulevard
Rockville, MD 20850

Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, April 5-9, 1988.

Special thanks is extended to Steven Ingels for his assistance in the presentation of this paper.

*** The research upon which this paper is based was sponsored by the Center for Education Statistics, U.S. Department of Education. Opinions expressed are those of the author and do not represent official positions of the Federal Government.**

THE NELS:88 SCHOOL SURVEY

Introduction

The National Education Longitudinal Study of 1988 (NELS:88) has become the third in a series of longitudinal studies sponsored by the Center for Education Statistics of the U.S. Department of Education. Building upon its predecessor studies, the National Longitudinal Study of the High School Class of 1972 (NLS-72) and High School and Beyond (HS&B), NELS:88 is designed to provide trend data about the critical transitions experienced by young people, as they develop, attend school, and embark on careers.

NELS:88 focuses on a sample of students who are enrolled in the eighth grade in the spring of 1988. Sample projections estimate that approximately 26,000 eighth grade students will be selected at random from a nationally representative sample of approximately 1,000 schools. The overall design for the study includes four major component surveys -- of students, parents, school administrators and teachers -- and a cognitive test component intended as a supplement to the student survey. In addition, the research design stipulates a base year data collection period during the spring of 1988, with follow-up surveys at two-year intervals which are intended to facilitate long-term trend analyses.

This paper is intended to: (a) identify the purpose and objectives of the School Survey component of NELS:88, (b) describe the survey administration procedures, and (c) present the results of the field test of the NELS:88 School Survey undertaken in the spring of 1987. Each of these items are addressed in the remaining sections of this paper following a brief overview of related longitudinal studies that have influenced the development and undertaking of NELS:88.

Background

This year, 1988, marks the twentieth anniversary since Congress has mandated that the U.S. Department of Education, Center for Education Statistics (CES) "... collect and disseminate statistics and other data related to education in the United States" (section 406(b) of the General Education Provisions Act, as amended by 20 U.S.C. 1221e-1). Since the passage of this mandate, educational policy-makers and researchers, at all levels, have advised CES to include in their response to this mandate the provision of data that would provide a better understanding

of the transitions made by young people from education to work, from dependence to independence, and from youth to adulthood.

A response to these data needs was initiated in 1970 with the establishment of the Longitudinal Studies Branch (LSB) of CES. The purpose of the LSB research program is to gather longitudinal data on nationally representative samples of students, which would aid in addressing such issues as the quality, equality, and diversity of educational opportunity and the effect of these factors upon individual development and educational and career outcomes. Since its creation, two major longitudinal studies have been launched and successfully undertaken. Included are: the National Longitudinal Study of the High School Class of 1972 (NLS-72), and High School and Beyond. Each of these programmatic efforts is discussed briefly below.

The National Longitudinal Study (NLS-72)

NLS-72 began in the spring of 1972 with a survey of a national probability sample of 19,001 seniors from 1,061 public, private and church-affiliated high schools. The sample was designed to be representative of the approximately three million high school seniors in more than 17,000 schools in the spring of 1972. Each sample member was asked to complete a student questionnaire and a 69-minute test battery. In addition, school administrators were asked to supply survey data on each student, as well as information about the school's programs, resources and grading system.

Five follow-ups conducted in 1973, 1974, 1976, 1979, and 1986 have been completed subsequently. At the time of the First Follow-up, an additional 4,450 students from the class of 1972 were added to the sample. Through intensive locating and tracking efforts, 83 percent of the 1972 Base Year sample and 85 percent of the expanded First Follow-up sample responded to the Fourth Follow-Up in 1979. Tracking efforts included the distribution of periodic newsletters to sample members in order to maintain their interest and cooperation and to update addresses for future contacts.

In addition to background information, the NLS-72 Base Year and Follow-up Surveys collected data on respondents' educational activities, such as schools attended, grades received, and degree of satisfaction with their educational institutions. Participants also were asked about work experiences, periods of unemployment, job satisfaction, military service, marital status, and children. Attitudinal information on self-concept, goals, participation in political activities, and ratings of their high schools are other topics for which respondents have supplied information.

High School and Beyond (HS&B)

The second major study undertaken by the CES was HS&B whose base year data collection efforts were initiated in the spring of 1980. HS&B was initiated in order to capture changes that have occurred in educational and social conditions, Federal and state programs, and needs and characteristics of students since the time of NLS-72. Thus, HS&B was designed to maintain the flow of relevant data to policy-makers at all levels who need to base their decisions on information that is reliable, relevant to the issues at hand, and current.

For the Base Year data collection, students were selected through a two-stage probability sample with schools as the first-stage units and students within schools as the second-stage units. There were 1,015 public, private and church-affiliated secondary schools in the sample and a total of 58,270 participating students. Unlike NLS-72, HS&B cohorts included both tenth graders and twelfth graders.

The base year survey included the administration of questionnaires and cognitive tests to over 30,000 sophomores and 28,000 seniors. Questionnaires focused on obtaining information on the educational and vocational choices made between the sophomore and senior years. In addition, the questionnaires covered school experiences, activities, plans, selected background characteristics, and language proficiency. By surveying sophomores again in their senior year, information was obtained on critical decisions made shortly before graduation, as well as information on high school dropouts. Unlike NLS-72, HS&B included a collection of data on the factors affecting family-formation behavior, intellectual development, and social participation.

Data were also collected from the school administrator in each participating school requesting information about school policies, programs, and student body characteristics. In addition, teachers provided comments on students in the sample. Finally, a sample of parents of sophomores and seniors was surveyed primarily for information about the financing of higher education.

Since the base year data collection in 1980, there have been three follow-ups of the HS&B cohorts, one in the spring of 1982, one in the spring of 1984, and the other in the spring of 1986. The third follow-up included the collection of financial aid record information for the 1980 high school seniors and sophomores who reported attending any postsecondary institution after

high school. For the 1980 sophomores, postsecondary transcripts were also included in this follow-up data collection.

The NELS:88 Research Design

NELS:88 seeks to gather trend data about critical transitions experienced by young people as they develop, attend school, and embark on careers. NELS:88 focuses on a sample of students enrolled in the eighth grade in the 1987-88 academic year. Similar to its predecessor studies, NELS:88 focuses on the collection of policy-relevant information related to such topics as effective schools, discipline, homework, course taking patterns, cognitive ability, dropouts, private schools, vocational education, special education, instruction for limited English speaking students, postsecondary access and choices, student financial assistance, employment stability, family formation, and graduate/professional training.

The base-year data collection period for NELS:88 began in February 1988 and will continue through June 1988. The sample of students will be administered a student questionnaire and a cognitive test battery intended to measure cognitive growth over time in reading, mathematics, science, and social studies. These data are to be supplemented by a school survey directed at the principals in schools where the participating students are enrolled, a survey of selected teachers of participating students, and a parent survey of a sample of participating students' parents. In addition, other components of the above surveys provide supplemental details concerning language minority students, gifted and talented students, and mathematics and science curriculum.

School Survey Design

The primary purpose of the School Survey is to gather general descriptive information about the educational setting or environment associated with the individual students who have been selected for participation in NELS:88. Information obtained through the survey is to be linked to each sampled student and is intended to meet the following objectives: (a) to assist in analyzing the learning environment and experiences of eighth grade students, and (b) to assist in distinguishing among different patterns of eighth grade schools and the effects of such patterns on the transitions of students to the tenth grade and beyond.

In keeping with the data needs targeted by NELS:88, emphasis is placed on gathering longitudinal data that would be useful in explaining future outcomes as assessed in follow-up survey efforts. Overall, the school description provided by the School Survey is intended to address the academic climate in terms of items such as enrollments and educational offerings as well as specific school policies. Thus, the following information needs were targeted for the School Survey:

- General school characteristics;
- Grading and/or testing structure used by the school;
- Program and facilities information;
- School climate;
- Parent interactions/involvement;
- Teaching staff characteristics; and
- Respondent data for follow-up.

Purpose and Content of the School Questionnaire

One of the initial tasks undertaken by project staff was the design and development of the school questionnaire. Initially, a content outline was prepared detailing possible information topics, suggesting item formats, and recommending specific items for inclusion in the questionnaires based on their use and performance in prior longitudinal studies (i.e., NLS-72 and HS&B). Following review and input by CES staff and members of a National Advisory Panel, the content outline was expanded into a draft questionnaire. After additional reviews and revisions as well as the development of supporting documentation, the school questionnaire was approved by FEDAC/OMB. Clearance was granted in December 1986 allowing the questionnaire to be formatted and printed in preparation for its administration during the Field Test.

The school questionnaire is designed to be responded to primarily by the school principal who is asked to provide descriptive information about the school's student population, teaching staff, policies, program offerings, and overall educational climate. In some cases, questionnaire items request factual data which may be provided by an individual who extracts such information from school files under the guidance of the school principal. Directions accompanying the questionnaire do indicate that portions of the questionnaire request factual data that may not be readily available from school records. In such cases, informed estimates are acceptable

responses to such questions. Finally, standard assurances of confidentiality and anonymity are provided.

The school questionnaire utilized in the field test contained eight parts. Each part is identified and described below.

- Part 1 - School Characteristics requests general information about the characteristics of the participating school in order to confirm its classification in the NELS:88 sample.
- Part 2 - Student Characteristics seeks information about the general characteristics of the school's student population.
- Part 3 - Teaching Staff Characteristics identifies general information about the professional teaching staff, related to group characteristics and operational organization within the school.
- Part 4 - School Policies and Practices requests information about the operational policies and practices of the school.
- Part 5 - Grading and/or Testing Structure seeks information about the grading and testing practices of the school.
- Part 6 - School Programs identifies information about the general academic and co-curricular programs that are available to eighth grade students enrolled in the school.
- Part 7 - School Climate identifies the overall morale and attitude of the students and professional staff regarding the school environment as perceived by the school principal.
- Part 8 - School-Parent Interactions collects information regarding the school's practices designed to involve parents in the educational process.

Respondent Sample for the School Questionnaire

Identification of the sample of respondents for the school questionnaire was quite straightforward, following the identification of schools involved in the Field Test. Each principal of the participating schools in the NELS:88 Field Test was requested to respond to the school questionnaire. In total, 51 school questionnaires were administered, and 46 completed questionnaires were returned, resulting in a 90 percent response rate. Of the remaining questionnaires, three principals (6%) refused to return the questionnaire, and two principals (4%) were categorized as non-respondents. Additional details, related to the school questionnaire response rates, are presented in a later section in this paper.

Data Collection Procedures

Following the preparation and printing of the approved school questionnaire and survey materials, as well as the identification of the sample of respondents, data collection procedures were initiated. Data collection includes all activities associated with obtaining completed questionnaires. These activities include: questionnaire mailout, receipt control, nonresponse follow-up, and data retrieval follow-up. Each activity is described in detail in the following sections.

Mailout. Distribution of the school and teacher questionnaires was accomplished in tandem by preparing a unique "school package" for each participating Field Test school. Each package contained the required number of teacher and school questionnaires, related cover letters, and instructional materials for the return of the school package.

The planned distribution method (method 1) involved transmitting a "school package" to the assigned school coordinator in order that the package would arrive at least two weeks prior to the scheduled survey day. This allowed time for the questionnaires to be distributed and completed before survey day. Thus, completed questionnaires could be collected by the team leader on survey day. During the Field Test, 23 schools were sent packages by this method.

For the remaining 28 schools, the materials required to prepare a school package for mailout were not received two weeks before survey day. For example, additional time may have been required to obtain school cooperation or a student roster that was needed for student sampling. Therefore, an alternative data collection method (method 2) was used with these schools. This second method involved sending a questionnaire package to the team leader before survey day. The team leader then gave the package to the school coordinator on survey day. The school coordinator was responsible for distributing and collecting questionnaires, and returning the completed forms.

Receipt Control. Records of all data collection activities were maintained in a computer-based receipt control system. This system was designed to maintain a database which included two files: a school file containing school-level information such as scheduled survey date and mailout date; and the questionnaire file containing information for each questionnaire, such as date received and status code.

When a "package" of questionnaires was returned, the contents were reviewed and status codes and dates were assigned and logged in the receipt control system. The questionnaires were then filed for further processing. The final status of questionnaire receipt by control of school (i.e., public, Catholic, and other private) appears in Table 1. In general, no trends in "refusals" or "nonresponse" were noted for the school questionnaire.

Table 1. School questionnaire status by control of school

School Questionnaire Status	Total	Control of School		
		Public	Catholic	Other Private
Nonresponse	2	2	-	-
Questionnaire received	46	34	5	7
Principal refused	3	2	1	-
Total	51	38	6	7

Nonresponse Follow-up. The three types of follow-up conducted for nonrespondent school questionnaires were: telephone calls to the coordinator when no package was received from a school, telephone calls to the coordinator when an incomplete package was received, and telephone calls to sampled respondents to collect questionnaire information by phone. Each of these types of follow-up is discussed in detail in the following sections.

After a questionnaire package was mailed, the expected return date was determined based on the mailout method used. An expected return date of one week after survey day was used for schools sent questionnaires by mailout method 1. For schools sent questionnaires by mailout method 2, the expected return date was three weeks after survey day.

Of the 51 schools in the Field Test, 38 returned packages on or before the expected return date. Seven schools returned packages within one week of the expected return date. For the remaining six schools, all of which used mailout method 2, prompting telephone calls were made to the coordinators. Five of the six schools responded after one telephone contact. The remaining school required two telephone contacts before responding.

The second type of nonresponse follow-up involved telephone calls to coordinators for incomplete packages. Sixteen incomplete packages were received during the Field Test. If

questionnaires were not received by the target date specified on the transmittal form, a follow-up telephone call was made to the school. Of the 16 incomplete packages, 11 required telephone calls to the coordinator.

For those questionnaires not received after follow-up telephone calls to the coordinator, the sampled respondents were contacted to collect questionnaire data by telephone. Five school questionnaires required this type of nonresponse follow-up. These contacts with respondents resulted in two refusals and three that would not complete questionnaires by phone but promised to mail them in. None of these forms was received, however.

Data Retrieval Follow-up. If a questionnaire contained a key question error or a data consistency error that related to a key question, a data retrieval telephone call was made to the respondent. When this occurred, all errors or missing data on the questionnaire were resolved. Table 2 shows the number of data retrieval follow-up cases by the control of the school (i.e., public, Catholic, and other private). No trends were noted in the degree of data retrieval follow-up by the control status of the Field Test school.

Table 2. Data Retrieval Followup Status by School Control

School questionnaire data retrieval status	Total	Control of School		
		Public	Catholic	Other Private
Data retrieval followup required	27	23	2	2
No data retrieval followup required	24	15	4	5
Total	51	38	6	7

All data retrieval follow-up telephone calls were conducted by telephone interviewers who were provided interactive training and a procedural manual. The manual included general instructions, administrative procedures, and instructions for handling specific data retrieval items.

Data Processing Completed Questionnaires

Data preparation and processing included all activities associated with converting information from completed questionnaires to clean data files. These activities included manual editing and coding, data entry, and machine editing, which are discussed in the following sections.

Manual Editing and Coding. This operation included manual editing and coding of the questionnaires. The purposes of the editing and coding operation were:

- To identify questionnaires needing data retrieval and prepare them for telephoning;
- To identify problem situations requiring coding decisions;
- To review completed data retrieval cases, recoding as necessary; and
- To prepare all questionnaires for data entry.

Data Entry. Coded questionnaires were transmitted in batches for data entry. Questionnaires were keyed to disk following specifications programmed specifically for the questionnaire. These specifications included all skip patterns and zero-filling of numeric fields. Each questionnaire batch was 100 percent key verified. After verification, each batch file was submitted for machine editing.

Machine Editing. The purpose of the machine edit operation was to identify and correct errors on the questionnaire data files. The types of errors corrected included respondent errors, coding errors, and data entry errors. All checks made during manual editing were also made during machine editing. Thus, any errors not identified manually were identified by computer. The machine edit specifications included the following types of checks: (a) alpha versus numeric checks, (b) range checks, (c) skip pattern checks, (d) data consistency checks, and (e) addition checks.

Analysis of Questionnaire Item Response and Nonresponse

As noted in a preceding section, the overall response rates achieved for the School Survey was 90 percent. Although this rate was quite acceptable, several suggestions and/or modifications were identified that would improve the quality of the information acquired through the survey instrument.

Overall, data collection and retrieval procedures related to the school questionnaire were judged to be successful. In total, 46 of the 51 questionnaires distributed were completed and returned. Of the remaining five questionnaires, three were considered refusals and two were categorized as nonresponses.

During the Field Test data collection stage, the return of "school packages" was monitored as was their completeness in terms of containing the expected number of school and teacher questionnaires. When a school package was delayed or determined to be incomplete, the school questionnaire was often the cause of the delay or incomplete status.

The delay in the return of the school questionnaire is reflective of the inordinate amount of time that it took to complete the questionnaire, as reported by school principals. As summarized from responses provided by the principals, the average number of person-hours required to complete the questionnaire was 1 hour and 36 minutes, with a range of 20 minutes to 5 hours. This time requirement certainly exceeded project staff projections and suggested that the overall length of the questionnaire be shortened.

The percentage of response to all items exceeded 90 percent at the conclusion of all follow-up efforts. Only a small number of items were responded to by principals as "don't know," or were coded by project staff as "not ascertained" in cases where non-key items were left blank by respondents. No items were coded as "refusals." It should be noted that these reported percentages are based on the number of expected responses for an item which varies due to item skip patterns.

As a result of coding and editing efforts as well as telephone data retrieval activities, several problem items were noted. Specific recommendations were suggested regarding the improvement of these items. Suggestions included: deleting response options, reformatting question stems and/or response options, or the deletion of an item.

Linkage of School Data to Student Files

One of the key objectives to be met by the School Survey is the linkage of descriptive information about schools to individual sampled students. The linkage of information about the sampled students' school environment is straightforward. That is, each sampled student's record is matched or linked to a single school record/questionnaire. Of the 1,556 eighth grade students,

descriptive school information was matched to a total of 1,396 individual students or 89.7 percent of the student sample. (Note: The total number of sampled eighth grade students varies with respect to the total number of eighth grade students presented in related reports or papers due to the varying number of eighth grade students sampled per school which was not adjusted for during the Field Test.) The 160 students (10.3 percent) who do not have matched school information are a result of the five school questionnaires coded as refusals or nonresponses.

In summary, the matching of school data to individual students was achieved for approximately 90 percent of the eighth grade student sample. Thus, no alterations were proposed to the School Survey design for the base year effort.

Summary of Recommendations

Overall, the Field Test of the School Survey was judged a success. Response rates achieved regarding the administration of the school questionnaire exceeded 90 percent. In addition, item response rates were more than sufficient, while respondents infrequently "refused" to respond to items or responded "Don't Know."

Although all Field Test activities were successfully conducted, several modifications to procedures, materials, and questionnaires were recommended. This final section, summarizes the recommendations that have been implemented in the undertaking of the NELS:88 Base Year study. Recommendations are presented and highlighted in the following three broad categories: pre-survey activities, data collection activities, and instrumentation.

Pre-survey activities. Prior to the distribution of questionnaires to respondents, several sources of information were required, and links between these sources had to be maintained in order to meet the overall study design goal of linking school information to individual sampled students. Some timing concerns were found in the acquisition and merging of data files and materials necessary for the administration of the school questionnaire resulting in the use of two mailing methods.

The alternative mailing plan (method 2) proved to be workable for the Field Test because of the small number of schools participating in the Field Test. The alternative mailing plan is being relied upon as little as possible for the base year study. When necessary, survey packages are express mailed to school coordinators to aid in their arrival prior to survey day. In

cases where this is not possible, the survey package is mailed to the school coordinator who is then responsible for the distribution, collection and return of questionnaires.

Data collection activities. Once questionnaire packages were transmitted, remaining survey activities were completed with very few problems. The added burden of tracking questionnaires through two different distribution methods is basically a logistical concern. Additional nonresponse efforts are required for missing and/or incomplete questionnaire packages returned by school coordinators (using the alternative distribution approach). The added logistical concerns and nonresponse follow-up efforts are hoped to be reduced during the base year effort by modifying timing of the collection of pre-survey information, as noted above.

Instrumentation. Directions for completing the questionnaire, individual wording of item stems, and item response scales were well understood by respondents and were effective in acquiring the targeted information. As reported, questionnaire and item nonresponse rates for the school questionnaire were quite low. However, the person-hours required by respondents to complete the questionnaire exceeded project staff projections. This finding made it necessary to reduce the length of the questionnaire to remain within the original response burden target. In addition, several of the specific modifications to problematic items were incorporated. Thus, the base year school questionnaire includes a total of 51 questionnaire items rather than 65, and are grouped into the following seven sections: school characteristics, student characteristics, teaching staff characteristics, school policies and practices, grading and/or testing structure, school programs and school climate.

THE NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988 TEACHER SURVEY*

Rocco P. Russo

Westat, Inc.
1650 Research Boulevard
Rockville, MD 20850

Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, April 5-9, 1988.

Special thanks is extended to David Bayless for his assistance in the presentation of this paper.

*** The research upon which this paper is based was sponsored by the Center for Education Statistics, U.S. Department of Education. Opinions expressed are those of the author and do not represent official positions of the Federal Government.**

THE NELS:88 TEACHER SURVEY

Introduction

The National Education Longitudinal Study of 1988 (NELS:88) has become the third in a series of longitudinal studies sponsored by the Center for Education Statistics of the U.S. Department of Education. Building upon its predecessor studies, the National Longitudinal Study of the High School Class of 1972 (NLS-72) and High School and Beyond (HS&B), NELS:88 is designed to provide trend data about the critical transitions experienced by young people as they develop, attend school, and embark on careers.

NELS:88 focuses on a sample of students who are enrolled in the eighth grade in the spring of 1988. Sample projections estimate that approximately 26,000 eighth grade students will be selected at random from a nationally representative sample of approximately 1,000 schools. The overall design for the study includes four major component surveys -- of students, parents, school administrators and teachers -- and a cognitive test component intended as a supplement to the student survey. In addition, the research design stipulates a base year data collection period during the spring of 1988, with follow-up surveys at two-year intervals which are intended to facilitate long-term trend analyses.

This paper is intended to: (a) identify the purpose and objectives of the Teacher Survey component of NELS:88, (b) describe the survey administration procedures, and (c) present the results of the field test of the NELS:88 Teacher Survey undertaken in the spring of 1987. Each of these items are addressed in the remaining sections of this paper following a brief overview of related longitudinal studies that have influenced the development and undertaking of NELS:88.

Background

This year, 1988, marks the twentieth anniversary since Congress has mandated that the U.S. Department of Education, Center for Education Statistics (CES) "... collect and disseminate statistics and other data related to education in the United States" (section 406(b) of the General Education Provisions Act, as amended by 20 U.S.C. 1221e-1). Since the passage of this mandate, educational policy-makers and researchers, at all levels, have advised CES to include

in their response to this mandate the provision of data that would provide a better understanding of the transitions made by young people from education to work, from dependence to independence, and from youth to adulthood.

A response to these data needs was initiated in 1970 with the establishment of the Longitudinal Studies Branch (LSB) of CES. The purpose of the LSB research program is to gather longitudinal data on nationally representative samples of students, which would aid in addressing such issues as the quality, equality, and diversity of educational opportunity and the effect of these factors upon individual development and educational and career outcomes. Since its creation, two major longitudinal studies have been launched and successfully undertaken. Included are: the National Longitudinal Study of the High School Class of 1972 (NLS-72), and High School and Beyond. Each of these programmatic efforts is discussed briefly below.

The National Longitudinal Study (NLS-72)

NLS-72 began in the spring of 1972 with a survey of a national probability sample of 19,001 seniors from 1,061 public, private and church-affiliated high schools. The sample was designed to be representative of the approximately three million high school seniors in more than 17,000 schools in the spring of 1972. Each sample member was asked to complete a student questionnaire and a 69-minute test battery. In addition, school administrators were asked to supply survey data on each student, as well as information about the school's programs, resources and grading system.

Five follow-ups conducted in 1973, 1974, 1976, 1979, and 1986 have been completed subsequently. At the time of the First Follow-up, an additional 4,450 students from the class of 1972 were added to the sample. Through intensive locating and tracking efforts, 83 percent of the 1972 Base Year sample and 85 percent of the expanded First Follow-up sample responded to the Fourth Follow-Up in 1979. Tracking efforts included the distribution of periodic newsletters to sample members in order to maintain their interest and cooperation and to update addresses for future contacts.

In addition to background information, the NLS-72 Base Year and Follow-up Surveys collected data on respondents' educational activities, such as schools attended, grades received,

and degree of satisfaction with their educational institutions. Participants also were asked about work experiences, periods of unemployment, job satisfaction, military service, marital status, and children. Attitudinal information on self-concept, goals, participation in political activities, and ratings of their high schools are other topics for which respondents have supplied information.

High School and Beyond (HS&B)

The second major study undertaken by the CES was HS&B whose base year data collection efforts were initiated in the spring of 1980. HS&B was initiated in order to capture changes that have occurred in educational and social conditions, Federal and state programs, and needs and characteristics of students since the time of NLS-72. Thus, HS&B was designed to maintain the flow of relevant data to policy-makers at all levels who need to base their decisions on information that is reliable, relevant to the issues at hand, and current.

For the Base Year data collection, students were selected through a two-stage probability sample with schools as the first-stage units and students within schools as the second-stage units. There were 1,015 public, private and church-affiliated secondary schools in the sample and a total of 58,270 participating students. Unlike NLS-72, HS&B cohorts included both tenth graders and twelfth graders.

The base year survey included the administration of questionnaires and cognitive tests to over 30,000 sophomores and 28,000 seniors. Questionnaires focused on obtaining information on the educational and vocational choices made between the sophomore and senior years. In addition, the questionnaires covered school experiences, activities, plans, selected background characteristics, and language proficiency. By surveying sophomores again in their senior year, information was obtained on critical decisions made shortly before graduation, as well as information on high school dropouts. Unlike NLS-72, HS&B included a collection of data on the factors affecting family-formation behavior, intellectual development, and social participation.

Data were also collected from the school administrator in each participating school requesting information about school policies, programs, and student body characteristics. In

addition, teachers provided comments on students in the sample. Finally, a sample of parents of sophomores and seniors was surveyed primarily for information about the financing of higher education.

Since the base year data collection in 1980, there have been three follow-ups of the HS&B cohorts, one in the spring of 1982, one in the spring of 1984, and the other in the spring of 1986. The third follow-up included the collection of financial aid record information for the 1980 high school seniors and sophomores who reported attending any postsecondary institution after high school. For the 1980 sophomores, postsecondary transcripts were also included in this follow-up data collection.

The NELS:88 Research Design

NELS:88 seeks to gather trend data about critical transitions experienced by young people as they develop, attend school, and embark on careers. NELS:88 focuses on a sample of students enrolled in the eighth grade in the 1987-88 academic year. Similar to its predecessor studies, NELS:88 focuses on the collection of policy-relevant information related to such topics as effective schools, discipline, homework, course taking patterns, cognitive ability, dropouts, private schools, vocational education, special education, instruction for limited English speaking students, postsecondary access and choices, student financial assistance, employment stability, family formation, and graduate/professional training.

The base-year data collection period for NELS:88 began in February 1988 and will continue through June 1988. The sample of students will be administered a student questionnaire and a cognitive test battery intended to measure cognitive growth over time in reading, mathematics, science, and social studies. These data are to be supplemented by a school survey directed at the principals in schools where the participating students are enrolled, a survey of selected teachers of participating students, and a parent survey of a sample of participating students' parents. In addition, other components of the above surveys provide supplemental details concerning language minority students, gifted and talented students, and mathematics and science curriculum.

Teacher Survey Design

The primary purpose of the Teacher Survey is to link information regarding specific teacher characteristics and practices to information about the characteristics and educational outcomes of the participating eighth grade students. The primary objective to be meet by the information obtained through this survey is to assist in acquiring an understanding of the effects of teaching on longitudinal student outcomes. Thus, data are collected that address teacher qualifications, specific course activities, curriculum exposure, and student-specific characteristics as judged by teachers. The overall design of the teacher survey reflects the four curriculum areas targeted by the overall goals established for NELS:88. These areas are: mathematics, science, English, and social studies.

The teacher questionnaire stresses the acquisition of information that may help account for the sampled students' subsequent development, particularly in the educational arena. Emphasis is placed on meeting the longitudinal student data needs targeted by NELS:88 as well as enabling cross-sectional analyses of students' educational attainments at the eighth grade level. Thus, the following information needs were identified:

- Teacher Impressions of Individual Students -- Specific information related to the teacher's impressions of school-related interests, aptitudes, and achievements of individual students.
- Curriculum Content -- Data about the name, ability track, curriculum, text, etc., for each course currently being taught by the teacher to one or more of the sampled students.
- Teacher Background and Activities -- General information about the teacher's demographics, qualifications and experience, current teaching status, duties, instructional methods, and influence over school and classroom policies.

Purpose and Content of the Teacher Questionnaire

One of the initial tasks undertaken by project staff was the design and development of the teacher questionnaire. Initially, a content outline was prepared detailing possible information topics, suggesting item formats, and recommending specific items for inclusion in the questionnaires based on their use and performance in prior longitudinal studies (i.e., NLS-72 and

HS&B). Following review and input by CES staff and members of a National Advisory Panel, the content outline was expanded into a draft questionnaire. After additional reviews and revisions as well as the development of supporting documentation, the teacher questionnaire was approved by FEDAC/OMB. Clearance was granted in December 1986 allowing the questionnaire to be formatted and printed in preparation for its administration during the Field Test.

The teacher questionnaire is designed to be responded to by teachers of the sampled eighth grade students. However, it is important to note that the teacher survey is not targeted at acquiring teacher-level data. Rather, it is targeted at identifying student-level data as reported by teachers, pertaining to specific eighth grade students and the mathematics, science, English, and social studies courses in which they are enrolled. Teachers are asked to respond to the questionnaire items relative to a specific list of sampled eighth grade students and to designated classes in which the sampled students are enrolled. The standard assurances of confidentiality and anonymity are also included.

The questionnaire contains three distinct sections:

- Part I - Student Information asks the identified teachers to indicate which of the sampled students they have had in their classes during the current academic year, and for those students enrolled in their class(es), to evaluate whether or not the student has had various school-related problems and handicaps.
- Part II - Class Information requires the teacher to respond to a series of course-related questions regarding a distinct set of classes they have been identified as teaching to one or more of the sampled students. Subsections of items within this segment of the questionnaire apply to the four specific curriculum areas (i.e., mathematics, science, English, and social studies), enabling teachers to respond to these subsections as appropriate.
- Part III - Teacher Background and Activities requests teachers to provide general background information about themselves and their school.

Respondent Sample for the Teacher Questionnaire

The sample design established for the identification of teachers is based on the need to obtain information about the NELS:88 sample students' school-related characteristics and experiences, especially those that might account for the students' educational development. The

teacher survey design is guided by the student sample and the four pre-established curriculum areas (i.e., mathematics, science, English, and social studies) that are the primary focus of NELS:88. At each school, approximately 30 students were sampled for participation in the Field Test. To achieve the objective of "linking information from the teacher to data about individual students in the NELS:88 sample," two teachers were identified as respondents to the teacher questionnaire for each student.

Selection of respondents to the teacher questionnaire for each student is based on the assignment of two curricula areas per school included in the Field Test sample. Following the listing of the sampled schools in order by state (i.e., CA, FL, IL, NY, and TX) within level within control; the random sequencing of the four curriculum area combinations (i.e., 1-Mathematics and English, 2-Science and Social Studies, 3-Mathematics and Social Studies, and 4-Science and English); and the selection of a random start combination (i.e., combination 3-Mathematics and Social Studies), curriculum areas were assigned to schools in repeating cycles of combinations, 3, 4, 1, 2. Thus, each Field Test school was randomly assigned one of the following combinations of curriculum areas:

- Mathematics and English;
- Mathematics and Social Studies;
- Science and English; or
- Science and Social Studies.

This selection procedure was designed to assure representation of Mathematics or Science curriculum areas and also of English or Social Studies in all participating schools. The curriculum area combinations of English and Social Studies as well as Science and Mathematics were not components of the design. In addition, the design achieved balanced representation of the four curriculum area combinations across the school variables of control (i.e., Catholic, Other Private, and Public) and level (i.e., Elementary, Middle, Junior-Senior High School). To a lesser degree, the design achieved balance across the five states covered by the Field Test sample.

Following the selection of the student sample in a given Field Test school and the random assignment of the curriculum combination areas, a matrix of information was obtained from school records. The Class Schedule matrix consisted of 32 rows (one per sampled student)

and two columns (one for each assigned curriculum area). For each cell in the matrix, that is, for each student-curriculum combination, the following information was recorded:

- Name of the student's current teacher in that subject;
- Weekly schedule for the class (e.g., period 3);
- Unique identifier for the student-teacher-subject combination (e.g., M-1 for the first student's current math class); this designation was used for all other sampled students who are in that same class-scheduled combination; and
- Course title.

Therefore, at any given school, each sampled student's current teacher in each of the two designated curriculum areas was selected to receive a teacher questionnaire. Results of the use of this sampling scheme indicated that the range of the number of teacher-respondents per sampled school was 1 to 15, with an average number of 5.9 teacher-respondents per school.

A total of 302 teachers in the 51 participating Field Test schools were requested to respond to the teacher questionnaire. Of these, 284 completed and returned questionnaires. Of the remainder, seven teachers (2%) refused to return the questionnaire; three teachers (1%) were categorized as nonrespondents; two teachers (1%) were no longer employed by the school; and six teachers (2%) were sampled in error. Thus, a response rate of 96 percent (284/302-6) was achieved. Additional details, related to the teacher questionnaire response rates, are presented in a later section of this paper.

Data Collection Procedures

Data collection activities were predicated on the completion and return of a Class Schedule matrix by the school coordinator. The Class Schedule fulfills two main purposes: (1) to identify, by teacher, the specific class each student attends for each of the curriculum areas assigned to the school, and (2) to link teacher ratings of students and descriptions of curriculum and practices to individual sampled students.

The Class Schedule must be received and processed before teacher questionnaires are sent to the school coordinator ideally, two weeks prior to survey day. The Class Schedule

information is used to produce, for each teacher-respondent, a list of classes for which the teacher is asked to provide descriptive information in part two of the questionnaire.

In general, the Class Schedule matrix and related materials worked quite well. In some cases, coordinators telephoned with clarification questions (10), and other coordinators made errors in completing the matrix which necessitated telephone follow-up contacts to clarify discrepancies (9). All cases involved minor questions or discrepancies that can be avoided during the base year effort by clarifying the instructions for completing the Class Schedule.

Coordinators were requested to complete and return a Class Schedule matrix following the selection of the student sample in each participating school. The completion and processing of the Class Schedule matrix was to be completed at least 10 to 12 working days prior to the Field Test school's assigned survey day. In more than one-half the cases, this target completion date was not met. As a result, in these cases an alternative plan was needed for the distribution and return of completed questionnaires.

Following the preparation and printing of the approved teacher questionnaire and survey materials, as well as the identification of the sample of teacher-respondents, data collection procedures were initiated. Data collection includes all activities associated with obtaining completed questionnaires. These activities include: questionnaire mailout, receipt control, nonresponse follow-up, and data retrieval follow-up. Each activity is described in detail in the following sections.

Mailout. Distribution of the teacher and school questionnaires was accomplished in tandem by preparing a unique "school package" for each participating Field Test school. Each package contained the required number of teacher and school questionnaires, related cover letters, and instructional materials for the return of the school package.

The planned distribution method (method 1) involved transmitting a "school package" to the assigned school coordinator in order that the package would arrive at least two weeks prior to the scheduled survey day. This allowed time for the questionnaires to be distributed and completed before survey day. Thus, completed questionnaires could be collected by the team leader on survey day. During the Field Test, 23 schools were sent packages by this method.

For the remaining 28 schools, the materials required to prepare a school package for mailout were not received two weeks before survey day. For example, additional time may have been required to obtain school cooperation, a student roster that was needed for student sampling, or a completed Class Schedule matrix. Therefore, an alternative data collection method (method 2) was used with these schools. This second method involved sending a questionnaire package to the team leader before survey day. The team leader then gave the package to the school coordinator on survey day. The school coordinator was responsible for distributing and collecting questionnaires, and returning the completed forms.

Receipt Control. Records of all data collection activities were maintained in a computer-based receipt control system. This system was designed to maintain a database which included two files: a school file containing school-level information such as scheduled survey date and mailout date; and the questionnaire file containing information for each questionnaire, such as date received and status code.

When a "package" of questionnaires was returned, the contents were reviewed and status codes and dates were assigned and logged in the receipt control system. The questionnaires were then filed for further processing. The final status of questionnaire receipt by control of school (i.e., public, Catholic, and other private) appears in Table 1. In general, no trends in "refusals" or "nonresponse" were noted for the teacher questionnaire.

Table 1. Teacher Questionnaire Status by Control of School

Teacher Questionnaire Status	Total	Control of School		
		Public	Catholic	Other Private
Nonresponse	3	3	-	-
Questionnaire received	284	254	11	19
Teacher refused	7	7	-	-
Teacher left school	2	2	-	-
Teacher sampled in error	6	5	-	1
Total	302	271	11	20

Nonresponse Follow-up. The three types of follow-up conducted for nonrespondent teacher questionnaires were: telephone calls to the coordinator when no package was received from a school, telephone calls to the coordinator when an incomplete package was received, and telephone calls to sampled respondents to collect questionnaire information by phone. Each of these types of follow-up is discussed in detail in the following sections.

After a questionnaire package was mailed, the expected return date was determined based on the mailout method used. An expected return date of one week after survey day was used for schools sent questionnaires by mailout method 1. For schools sent questionnaires by mailout method 2, the expected return date was three weeks after survey day.

Of the 51 schools in the Field Test, 38 returned packages on or before the expected return date. Seven schools returned packages within one week of the expected return date. For the remaining six schools, all of which used mailout method 2, prompting telephone calls were made to the coordinators. Five of the six schools responded after one telephone contact. The remaining school required two telephone contacts before responding.

The second type of nonresponse follow-up involved telephone calls to coordinators for incomplete packages. Sixteen incomplete packages were received during the Field Test. If questionnaires were not received by the target date specified on the transmittal form, a follow-up telephone call was made to the school. Of the 16 incomplete packages, 11 required telephone calls to the coordinator.

For those questionnaires not received after follow-up telephone calls to the coordinator, the sampled respondents were contacted to collect questionnaire data by telephone. Nine teacher questionnaires required this type of nonresponse follow-up. These contacts with respondents resulted in seven refusals and two that would not complete questionnaires by phone but promised to mail them in. None of these forms was received, however.

Data Retrieval Follow-up. If a questionnaire contained a key question error or a data consistency error that related to a key question, a data retrieval telephone call was made to the respondent. When this occurred, all errors or missing data on the questionnaire were resolved.

Table 2 shows the number of data retrieval follow-up cases by the control of the school (i.e., public, Catholic, and other private). No trends were noted in the degree of data retrieval follow-up by the control status of the Field Test school. However, a large proportion of teacher questionnaires required data retrieval follow-up. Additional details, related to the data retrieval follow-up effort, are presented in a later section of this paper.

Table 2. Data Retrieval Follow-up Status by School Control

Teacher questionnaire data retrieval status	Total	Control of School		
		Public	Catholic	Other Private
Data retrieval follow-up required.....	225	201	9	15
No data retrieval follow-up required.....	77	70	2	5
Total.....	302	271	11	20

All data retrieval follow-up telephone calls were conducted by telephone interviewers who were provided interactive training and a procedural manual. The manual included general instructions, administrative procedures, and instructions for handling specific data retrieval items.

Data Processing Completed Questionnaires. Data preparation and processing included all activities associated with converting information from completed questionnaires to clean data files. These activities included manual editing and coding, data entry, and machine editing, which are discussed in the following sections.

Manual Editing and Coding. This operation included manual editing and coding of the questionnaires. The purposes of the editing and coding operation were:

- To identify questionnaires needing data retrieval and prepare them for telephoning;
- To identify problem situations requiring coding decisions;

- To review completed data retrieval cases, recoding as necessary; and
- To prepare all questionnaires for data entry.

Data Entry. Coded questionnaires were transmitted in batches for data entry. Questionnaires were keyed to disk following specifications programmed specifically for the questionnaire. These specifications included all skip patterns and zero-filling of numeric fields. Each questionnaire batch was 100 percent key verified. After verification, each batch file was submitted for machine editing.

Machine Editing. The purpose of the machine edit operation was to identify and correct errors on the questionnaire data files. The types of errors corrected included respondent errors, coding errors, and data entry errors. All checks made during manual editing were also made during machine editing. Thus, any errors not identified manually were identified by computer. The machine edit specifications included the following types of checks: (a) alpha versus numeric checks, (b) range checks, (c) skip pattern checks, (d) data consistency checks, and (e) addition checks.

Analysis of Questionnaire Item Response and Nonresponse

As noted in a preceding section, the overall response rate achieved for the Teacher Survey was 96 percent. Although this rate was quite acceptable, several suggestions and/or modifications were identified that would improve the quality of the information acquired through the survey instruments.

Overall, data collection and retrieval procedures related to the teacher questionnaire were judged to be successful. In total, 284 of the 302 questionnaires distributed were completed and returned. Of the remaining 18 questionnaires, seven were considered refusals, three were categorized as nonresponses, two were not returned because the teacher-respondent left the sampled school, and six teacher-respondents were sampled in error by the school coordinators.

As previously discussed, approximately 5.9 teacher respondents were identified per Field Test school. On the average, each teacher provided ratings for 10.5 students in Part I of the

questionnaire. Only 15 percent of the teacher-respondents were required to provide ratings for more than 18 students. However, six teacher-respondents were required to provide ratings for all 32 sampled students in their school.

In Part II of the questionnaire, an average of 2.8 classes were described by each teacher-respondent. In total, 790 teacher-class ratings were provided by the 284 teacher-respondents. Of this total, 208 were identified as English teacher-class combinations, 207 were identified as science teacher-class combinations, 191 were identified as social studies teacher-class combinations, and 184 were identified as mathematics teacher-class combinations. Table 3 presents the frequency distribution of the number of classes rated by the teacher-respondents. In total, 12 teacher-respondents (4.3 percent) were required to rate more than five classes.

Table 3. Number of Courses Rated by Teacher-Respondents

Number of Courses	Number of Respondents	Percent
1	73	25.7
2	69	24.3
3	50	17.6
4	46	16.2
5	34	11.9
6	11	3.9
9	1	0.4
	284	100.0

In general, the percentage of response to all items equaled or exceeded 90 percent. Only, ten items yielded lower response rates, ranging from 83 percent to 88 percent. Teacher-respondents infrequently provided a "Don't Know" or "Refused" response to items. It should be noted that these reported percentages are based on the number of expected responses for an item which varies due to item skip patterns.

Data retrieval follow-up efforts were required for 225 questionnaires or 74.5 percent. In part, this high percentage of follow-up effort was due to the amount of time that it took to complete the questionnaire, as reported by teacher-respondents. The average number of person-hours required to complete the questionnaire was 1 hour and 9 minutes, with a range of 10 minutes to 5 hours. This time requirement was 15 percent over the one hour response burden

target established for the teacher questionnaire and suggests that the overall length of the questionnaire be shortened.

A second factor relative to the high follow-up effort relates to discrepancies noted during the editing of return questionnaires. The editing process discovered apparent confusions in responses indicating that a teacher-respondent did not have one or more of the sampled students assigned to his/her class while the Class Schedule matrix completed by the coordinator indicated that the sampled student was assigned to the teacher's class. Follow-up calls determined that most discrepancies were the result of errors made by the teachers. In some cases, teachers interpreted the Part I questionnaire items to include only students assigned to them continuously since September, instead of at any time since September. While in other cases, teachers scanned the list of student names quickly and skipped some students they should have included.

Finally, as a result of coding and editing efforts as well as telephone follow-up activities, several problem items were noted. Specific recommendations were suggested regarding the improvement of these items. Suggestions included: deleting response options, reformatting question stem and/or response options, or deletion of the item.

Linkage of Teacher Data to Student Files

One of the key objectives to be met by the Teacher Survey is the linkage of information about schools, teachers, curriculum, and teacher practices to individual sampled students. The linkage of data acquired through the teacher questionnaire is quite complex and is achieved solely through the Class Schedule matrix. The Class Schedule identifies, for each sampled student, the student's teachers in two pre-assigned curriculum areas and the specific courses (i.e., course title and class period or section number) in which the student is enrolled. Thus, the Class Schedule provides the mechanism by which teacher ratings of individual students can be matched to the sampled students as well as matching descriptive course information about the students' enrolled classes.

Data processing of the completed Class Schedules identified 3,112 possible student-teacher and student-course combinations. That is, two teachers and courses were identified for each of the 1,556 sampled eighth grade students. (Note: The total number of sampled eighth grade

students varies with respect to the total number of eighth grade students presented in related reports or papers due to the varying number of eighth grade students sampled per school, which was not adjusted for during the Field Test.)

Of the 3,112 student-subject combinations, individual student ratings and course information were matched to 2,845 student-subject combinations or 91.4 percent of the expected combinations. Of the remaining 267 combinations, 74 were not matched to teacher ratings due to student withdrawals from school, three students were not enrolled in courses in the pre-assigned curriculum areas, and 190 combinations were not matched due to missing teacher questionnaires (144) or due to a teacher-respondent's indicating that he/she did not have the sampled student enrolled in his or her class (46).

Although each sampled student was expected to be matched to two teacher ratings, 142 students were provided ratings by more teacher-respondents than was expected. This additional yield of student-teacher information is a positive, yet unpredictable, result of the Teacher Survey design.

In summary, the matching of teacher data to individual students was achieved for approximately 91 percent of the eighth grade student sample. Thus, no alterations are proposed to the Teacher Survey designs for the base year effort.

Summary of Recommendations

Overall, the Field Test of the Teacher Survey was judged a success. The overall questionnaire response rate achieved was 96 percent. In addition, item response rates were more than sufficient, while respondents infrequently "refused" to respond to items or responded "Don't Know."

Although all Field Test activities were successfully conducted, several modifications to procedures, materials, and questionnaires have been suggested. This final section, summarizes the recommendations that have been implemented in the undertaking of the NELS:88 Base Year study. Recommendations are presented and highlighted in the following three broad categories: pre-survey activities, data collection activities, and instrumentation.

Pre-survey activities. Prior to the distribution of questionnaires to respondents, several sources of information were required, and links between these sources had to be maintained in order to meet the overall study design goal of linking teacher information to individual sampled students. Some timing concerns were found in the acquisition and merging of data files and materials necessary for the administration of the teacher questionnaire resulting in the use of an alternative mailing method.

The alternative mailing plan (method 2) proved to be workable for the Field Test because of the small number of participating schools. However, use of this alternative mailing method is to be relied upon as little as possible during the base year effort. In order to ensure that Class Schedule forms are completed in a timely manner, and that "school packages" arrive prior to the school's survey day, these materials are transmitted via express mail, when necessary.

In cases where the "school packages" is not transmitted prior to survey day, it is mailed directly to the coordinator who is then responsible for the distribution, collection and return of questionnaires.

Data collection activities. Once questionnaire packages were transmitted, remaining survey activities were completed with very few problems. The added burden of tracking questionnaires through two different distribution methods is basically a logistical concern. Additional nonresponse efforts were required for missing and/or incomplete questionnaire packages returned by school coordinators (using the alternative distribution approach). The added logistical concerns and nonresponse follow-up efforts are hoped to be reduced during the base year effort by modifying timing of the collection of pre-survey information, as discussed above.

Instrumentation. Directions for completing the questionnaire, individual wording of item stems, and item response scales were well understood by respondents and were effective in acquiring the targeted information. As reported, questionnaire and item nonresponse rates for the school questionnaire was quite low. However, the person-hours required by respondents to complete the questionnaire exceeded project staff projections. This latter finding would suggest that the length of the questionnaire be reduced to remain within the original response burden targets. A review of data items and information needs basically resulted in very few changes to the questionnaire for the base-year study. Several of the specific modifications to problematic items

were incorporated. Particular attention was given to the clarification of the student rating segment of the questionnaire (Part I) in order to reduce potential discrepancies between teacher responses and the completed Class Schedule matrix. The modified teacher questionnaire contains the same three segments and generally, the same number of items.

The Design and Validation of the NELS Test Battery

by

Donald A. Rock

and

Judith M. Pollack

Educational Testing Service
Princeton, NJ 08541-0001

Paper presented at the American Educational Research Association Annual Meeting, New Orleans, Louisiana, April 1988.

The Design and Validation of the NELS Test Battery

Donald A. Rock & Judith M. Pollack

Educational Testing Service

The Educational Testing Service (ETS) was contracted by OERI to develop the National Educational Longitudinal Study (NELS) test battery to assess students' cognitive growth during their transition from grade eight to grade twelve. The students are to be tested in the spring of 1988 (eighth grade), then retested in the spring of 1990 (tenth grade), and then retested again in the spring of 1992 as Seniors. The goals of the NELS longitudinal study are to estimate the extent and correlates of cognitive growth in four cognitive areas. The four cognitive areas are:

- o Reading or literacy
- o Mathematics
- o Science
- o History and Government

Since the content coverage is quite broad given the testing time allotted, (one hour and twenty five minutes) special testing procedures needed to be implemented. In addition the tests had to be constructed to reliably measure gain across a four year period. In order to "tailor" the tests to the abilities of the students at each grade level it was anticipated that three test forms would be constructed, one for each grade level within each content area. Item response theory (IRT) methodology could then be used, if the dimensionality assumptions are met, to put the grade level forms on the same scale. This grade level adaptive procedure was necessary in order to minimize floor and ceiling effects that typically distort estimates of gain scores. Item pools in all four content areas have been pretested and analyzed at each of the three grade levels. This paper will report the results of the pretest including the psychometric properties of the proposed test forms along with the patterns of gain that were observed across the three grade levels.

In addition to measuring growth and its educational antecedents, a second purpose of the NELS test battery is to provide a common scale such that comparable HS&B cohorts can be compared on a cross-sectional basis with the NELS cohort in the area of mathematics and possibly reading. Using sophomore data from the 1980 HS&B cohort and the comparable NELS sophomores in 1990, one can examine the change in both reading and mathematics performance that will have taken place during the eighties. The same comparison can be carried out for HS&B seniors in 1982 versus the NELS 1992 seniors.

In terms of longitudinal academic growth, the HS&B 1980 to 1982 Sophomore to Senior gains in mathematics and reading can be contrasted with

the comparable NELS cohort gains taking place between the 1990 sophomores and the 1992 seniors. The fact that the mathematics and reading tests are designed to be on the same scale allows for contrasts in both longitudinal trends as well as cross-sectional trends.

In order to better understand the NELS longitudinal testing design, it would be useful to contrast the goals and procedures of a large scale cross-sectional assessment design such as NAEP with those of a longitudinal design such as NELS. Both test batteries (NAEP and NELS) are subject to relatively severe limitations with respect to the amount of testing time that is available. Within these testing time limitations NAEP attempts to assess performance in a number of content areas at both the total score and at the subscore level when possible.

NAEP can be thought of as an assessment procedure analogous to a thermometer which takes the Nation's educational temperature in a number of content areas and at various stages of student development. Because of testing time limitations and the fact that NAEP will be attempting to develop subscores for various cross-classifications of content by cognitive process, the measurement reliability of these subscores will be relatively low. However, they will still be sufficiently reliable for computing scores such as means aggregated across the total sample receiving a particular subset of the measures.

Students in the NAEP administration receive only a subset of the measures in order to maximize the number of content areas that are covered in a specific administration. Since the particular block of measures a student receives is assigned in a random fashion, the aggregate (mean) scores based on this particular block of items and students can be assumed to be an unbiased estimate of the total student population's performance. This balanced incomplete block testing design enables the NAEP researchers to inform policy makers about the <u>average</u> performance of the nation's students in a particular content area. However, since a relatively small proportion of the total NAEP sample receives a given block or content area, estimates of student performance for finer cross-classifications of students can be relatively unstable. Cross-sectional trends in averages can also be estimated but the policy interpretations are subject to all the usual qualifications about using cross-sectional data to draw longitudinal inferences.

A longitudinal study such as NELS attempts to inform policy makers about questions having to do with how much change occurred and <u>why</u> students have changed or gained in performance over time. The emphasis in NELS is on both group and individual differences with respect to status at a particular point in time as well as to changes over time. In order to deal with policy questions that deal with both individual as well as group changes, longitudinal studies need to be somewhat more "high fidelity" than would be necessary for cross-sectional type analysis of aggregated data, given the same amount of testing time. High fidelity as used here simply means measuring fewer things with greater accuracy. Lower fidelity designs such as NAEP measure many things with somewhat less reliability. This lower fidelity design is appropriate for NAEP's primary purpose-the reporting of

aggregate scores without too fine a breakdown. Obviously if testing time were not limited (as are budgets and the patience of the nation's school systems) these compromises would not be necessary.

Longitudinal studies of gain require more reliable measurement since they attempt to correlate differences in school processes at the individual level with similar differences in gains in achievement at the individual level. One attempts to correlate differences between the educational processes received by Jane and Johnny with differences in their relative amounts of achievement gains.

In summary the balanced incomplete block design as used in NAEP is less appropriate for NELS type "individual" difference models. Having large samples receive more items maintains sufficient accuracy at the individual score level in order to:

- o reliably estimate correlates of achievement at a given time,
- o estimate extent and correlates of gain over time,
- o allow for finer cross-classifications in order to investigate whether the educational process works the same for different kinds of individuals.

Pretest Sample Design

The pretest sample included 1,376 students at 51 schools with an 8th grade; 1,131 students at 30 schools with a tenth grade, and 1,017 students at 30 schools with a 12th grade. The sample was drawn from five states: New York, California, Florida, Illinois, and Texas.

Two pretest forms were constructed-Form A and Form B. Form A included the pretest items for the Reading and the Science test. Form B included the Mathematics items and the Social Studies items. Table 1 below shows the number of pretest items in each content area and the test administration times.

Table 1

Numbers of Pretest Items and the Test Administration Times

Cognitive Area	Number of Items	Admin. Time
Reading (Form A)	50	55
Science (Form A)	42	30
Mathematics (Form B)	82	60
Social Studies (Form B)	50	25

The two test forms were spiralled within each grade level. That is, half of the students at each grade level received form A and the other half received form B. Since tests were nested under forms, this amounted to spiralling tests. A total of 1599 students across all three grade levels took Form A (Reading and Science). The Form A totals included 700 eighth graders, 470 tenth graders, and 429 twelfth graders. Similarly 1812 students took Form B (Mathematics and Social Studies). The Form B totals included 667 eighth graders, 599 tenth graders, and 546 twelfth graders. While these numbers fell somewhat short of the original target numbers, i.e., 600 on each form within each grade, it was felt that the sample size was still sufficient to estimate the necessary item parameters and select the items to be included in the final forms.

Table 2 below shows the percent of examinees that completed the tests by grade level.

Table 2

Percent of Examinees Completing each Test by Grade Level

Test Content	Grade 8	Grade 10	Grade 12
Social Stud.	97	97	98
Mathematics	87	92	89
Reading	98	99	98
Science	95	95	94

The percent of the pretest sample completing each test ranged from 87% of the students in the eighth grade on the Mathematics test to 98% of the students in the 12th grade on the Reading test. A test is considered relatively unspeeded if approximately 90% percent of the examinees complete the test. While the Mathematics did not quite achieve the 90% level of completion at the 8th and 12th grade, it was felt that it was close enough to be considered unspeeded. The number of items to be included in the projected final forms was estimated from these pretest timings.

Selecting Items For Measuring Change

The Technical Advisory Committee listed a number of concerns that they felt needed to be addressed in the test construction phase. The concerns were that the tests:

(1) Provide curriculum relevant content coverage.

(2) Provide scores in the four content areas.

(3) Should minimize floor and ceiling effects.

(4) Be equatable to the HS&B tests in the comparable areas.

(5) Emphasize "higher order thinking skills".

(6) Reliably measure both change and status at a given point in time.

(7) Provide reliable subscores if possible.

(8) Minimize the inclusion of items that were differentially difficult for sex or ethnic groups.

(9) Be relatively short so as not to unnecessarily burden the school systems.

Many of these suggestions are in conflict and some accommodation had to take place before the final forms could be arrived at. The next section discusses some of the compromises that were made.

Content Coverage

Test construction committees consisting of curriculum experts and teachers at the appropriate grade levels selected items from the NAEP, IEA, and HS&B item pools that were consistent with the curriculum objectives of the course work in the four content areas and three grades. The test constructors were instructed to emphasize higher order thinking skills in their item selection when such selection was not in opposition to curriculum relevance. These committee selections led to the pretest item pool.

The coverage of final item selections depended, of course, on how representative the original pretest item pool was with respect to curriculum objectives.

Minimizing Floor and Ceiling Effects

In order to arrive at efficient estimates of change scores, the final forms were constructed to minimize "floor" and "ceiling" effects. Floor effects occur when test forms are much too difficult for a large number of the examinees. When this occurs the low ability examinees receive chance scores on the first testing. Thus there is no discrimination among the low scoring individuals at the time one testing. If the test is still too difficult at the time of the second testing the same individuals will once again receive chance scores. In this case there will be no discrimination among low ability individuals at time one or time two and more importantly there will be no reliable discrimination among examinees with respect to amounts of change. Estimates of mean change for these individuals will also be inaccurate. If the test only has floor effects at time one for low ability individuals the situation is less serious. When this occurs the reliability of the <u>individual</u> score changes may be attenuated while mean changes for these individuals will be somewhat biased downward.

Ceiling effects can even be more damaging with respect to estimating gains accurately. If the test tends to have insufficient numbers of the more difficult items on the second and third testings, the change scores will be artificially constrained for high ability examinees. That is, the highest scoring individuals at time one are already at or close to the maximum possible score; if the maximum score is not "stretched" at time two by adding more difficult items, these examinees will appear to have gained very little. This phenomenon contributes to the typical finding of a negative correlation between "raw" gain scores and the initial test scores.

Given rather limited amounts of testing time a compromise solution was to develop test forms that were "adaptive" to the ability level of the examinee. That is, the examinee is presented items that are appropriate for his or her ability level. This procedure tends to minimize the possibility of floor and ceiling effects biasing the estimates of the score gains. A completely tailored test is best administered and scored on a computer. The level of "tailoring" of the item to the ability of the examinee can vary from computer based completely adaptive tests with all examinees receiving different items to mildly adaptive test forms appropriate for paper and pencil.

Our approach here was to build grade adaptive tests where necessary. Grade adaptive tests, i.e., tests whose distributions of item difficulties, vary by grade level rather than at the individual level, can be administered in a paper and pencil format under standardized conditions. Grade adaptive tests should be sufficiently adaptive to minimize floor and ceiling effects and yet maintain the same standardization of administration procedures as HS&B and NAEP. This will facilitate cross-walking scores between the various populations where appropriate.

There is no "free lunch" however; items used in adaptive tests must be scaleable which in turn means that they should form a reasonably unidimensional item pool. This allows one to put test scores based on different items that are given in the different grades on the same scale. Originally it was proposed that Mathematics, Reading, and Science be grade level adaptive. The requirement that the Social Studies test cover the domains of American History, Government, and Geography made it a less likely candidate for making it adaptive. Similarly the Science test specifications required items dealing with concepts in earth science, life science, chemistry and physics. In addition the Science items were written to emphasize higher order thinking skills in each of these four knowledge domains.

Not unexpectedly the Science and Social Studies test item pools turned out to be multidimensional. In the case of both the Science and Social Studies tests, they were designed by the test developers to deal with concepts which are familiar at all three grade levels thus somewhat mitigating the need to make them grade level adaptive. While the Reading and Mathematics test item pools were also designed to measure higher order thinking skills in a number of content areas, past experience suggests that their respective item pools would turn out to be reasonably unidimensional

and thus amenable to building adaptive tests. Factor analytic results of the mathematics and reading item pools supported the unidimensionality of these two domains. Given the unidimensionality of the Mathematics and Reading pools, it was decided to develop grade adaptive tests in both these areas. Conversely because of the multidimensionality and the potential familiarity at each grade level of the concepts in the Science test, one form was designed for all grade levels. Similar arguments led to the construction of one Social Studies form for all grade levels.

Modeling Change

One potential drawback of longitudinal designs that operate within a limited amount of testing time is that it is difficult to estimate subscores within a given test that have sufficient reliability for individual difference models including change measurement. Rather than give up the notion of having interpretable subscores in a longitudinal design, we propose a slightly different approach to calibrating and scoring the tests that will yield subscores with sufficient reliability for change measurement given certain assumptions of unidimensionality.

Basically this approach assumes that within the total test there are clusters of items (testlets) that are internally homogeneous with respect to difficulty level and that form a complexity hierarchy that in turn can be used to behaviorally anchor the total score scale. In order to behaviorally anchor the scale these testlets should form an interpretable hierarchy of skill levels. For example, in Mathematics the lowest skill level might be a testlet of items describing single operation arithmetic on whole numbers. The next skill level could be a cluster of items describing arithmetic problems requiring multiple operations etc. This approach is similar to that used in developing the NAEP behaviorally anchored reading scale except that we prefer to use clusters of items to mark a scale point because of the enhanced reliability of clusters over single items. These clusters that form the complexity hierarchies are consistent with the notion of an individual difference model that assumes that individuals are all at different points in their skill development. Any reasonable model of change should be able to pinpoint which of the complexity levels is showing the most change for a given individual. Changes in the total score over time, whether the score is in raw or residualized form, tell us little about where (at what level of complexity) the change is taking place unless the total score is behaviorally anchored.

The above behaviorally anchored scaling procedure (Rock & Pollack, 1987) has been developed and applied to the HS&B mathematics items and has demonstrated significant improvement over the traditional measures of change. In the case of the HS&B mathematics items the ordered scales in terms of complexity provided diagnostic information on the skill levels that showed the greatest changes as well as demonstrating significantly higher relationships between change and school processes.

Final Selection of Items

An item was selected for the final form if it met a number of substantive and psychometric criteria. Items were plotted in a series of two dimensional tables. The vertical axis was defined by gain over time: either the difference between 10th and 8th grade item difficulties or the difference between 12th and 10th grade item difficulties (i.e., proportion passing). The horizontal axis was defined by the level of the item- total biserials. Items were then selected when possible from the upper right hand quadrant with the restriction that they meet requirements with respect to coverage of objectives and also lead to a reasonable distribution of difficulties within a grade level form. The items in the upper right hand quadrant are those items that appear to be most sensitive to change on a cross-sectional basis and yet are more reliable. For example, an item would be a candidate for the mathematics 8th and 10th grade forms if it had a relatively high biserial (>.40) and a large increment in the proportion passing when going from the 8th to the 10th grade. Similarly, items whose p-values (difficulties) only discriminate between 10th and 12th graders might appear only in both the 10th and 12th grade forms.

In addition it is anticipated that there will be a subset of at least 15 to 20 items that will be common to all the mathematics grade forms. Similarly a lesser number will be common to all levels of the reading test. The majority of these common items will be from the comparable HS&B test. These common items will serve two purposes. They will provide a check on the vertical IRT equating of the reading and mathematics scores across grades as well as provide a check on the "bridge" with HS&B. It should be kept in mind that while forms have been proposed for all grade levels, only the 8th grade form is fixed in "concrete" and is presently being administered in the field. As indicated above the Science and Social Studies tests are projected to be the same for all grade levels.

Characteristics of the Proposed Final Forms

Reading

The proposed reading tests are grade level adaptive and include passages based on scientific material, social studies, and literature at all three grade levels. Virtually all passages were expository and require either interpretive or evaluation skills on the part of the reader. That is, the items attempt to test the reader's ability to divine the author's main idea, purpose, underlying assumptions, or motives. Others deal with extension of ideas or relationship between ideas. When the reading data was analyzed a single factor model fit quite well at all three grades. There are 21 items at grade 8 and 20 items in the remaining two grade forms.

Mathematics

The content areas covered in the mathematics tests (40 items) include arithmetic, algebra, data interpretation, geometry, and higher order

thinking skills (hots). The processes being tapped include knowledges and skills, understanding/comprehension, and problem solving. Since the mathematics tests are grade level adaptive, there are proportionately more arithmetic items in the 8th grade form than in the 10th or 12th grade forms. Similarly there is an increase in the number of geometry and high level problem solving items as one goes from the 10th to 12th grade. As in the case of the reading test a single factor model provided a good fit to the mathematics data at all three grade levels. A large first factor is, of course, necessary for the IRT vertical equating for the grade level adaptive approach taken here. Table 3 presents the content distribution of the mathematics test forms.

Table 3

Number of Items

	Grade 8	Grade10	Grade 12
Arithmetic	18	16	12
Algebra	9	10	8
Data Representation	4	5	3
Geometry	5	5	10
HOTS	4	4	7

Social Studies

The Social Studies test (30 items) content areas were about equally split between American History (13 items)and citizenship/government (14 items). There are also three geography items. Factor analysis was unable to retrieve these two content areas as discriminable factors. The same form is projected for all grade levels.

Science

The Science test (25 items) is about equally split in content between life science (9 items), earth science (9 items), and chemistry (7 items). The test construction emphasis was on concepts that would apply to all grade levels. As in the case of the Social Studies test the factor analysis solutions did not reflect the item classifications.

Psychometric Characteristics of the Projected Test Forms

The following psychometric descriptions refer to the final forms of the reading and history/government tests. The psychometric descriptions for the mathematics form at the 8th grade level is not exact since five of the items

in this revised form either had either no pretest information (2 items) or only pretest information from HS&B. Performance on these items was estimated for each individual in order to obtain the summary statistics. Similar cautions apply to the science test at all three grades.

Reading

Table 4 presents the summary statistics for the proposed reading test forms which would be administered at grades 8, 10, and 12. For each grade level, means, standard deviations, and reliabilities are presented for the total pretest item pool and the proposed final forms. On the right side of each table is the difference between the white group's mean and the selected ethnic group means in terms of standard deviation units.

****
Insert Table 4 about here
****

The coefficient Alpha estimates of internal consistency reliability of .83 for grade 8, and .85 for grades 10 and 12 are excellent for tests of this length. It is not appropriate to contrast the grade level raw score means at this time because of the different items involved.

Contrasts between the various ethnic groups and whites as shown by the effect size columns remain quite constant whether the full pretest item pool or the selected items are used. While the reliability for the Blacks is somewhat lower than that of the whites at both grade 8 and grade 10, the standard errors of measurement are virtually the same. The principal revision made to the Reading test since the original proposed test (Rock & Pollack, 1987) was to add additional easy items to the 8th grade form. This increased the reliability coefficient for Blacks and Hispanics from .77 to .80 (Blacks) and .78 to .80 (Hispanics). The revised 8th grade test had a mean difficulty of .593 which is very close to a theoretical optimum (.55 - .65). There does appear, however, to be a slight floor effect for Blacks at the 8th and to a lesser extent the 10th grade. It is our opinion that this effect is not of sufficient size to be of any great concern. It should be kept in mind that when measuring change the effect of greatest concern is the minimizing of potential ceiling effects.

The selected items for the 10th grade form also included 7 items from the HS&B Reading Test which should be sufficient to put the NELS items on the same scale as HS&B if one wishes to compare performance between the 1980 HS&B sophomore cohort and the comparable 1990 NELS sophomores.

Mathematics

The group statistics and estimated reliabilities of the 8th, 10th and the 12th grade forms are presented in Table 5.

****
Insert Table 5 about here
****

Inspection of the reliabilities across the three grades suggests that the proposed 40 item Mathematics test forms lost little or no reliability when the item pool was cut in half. The grade 8 forty item form maintained the full reliability of the total 82 item pool at the eighth grade and only dropped from .94 to .92 at the tenth grade. The forty item grade 12 form also demonstrated excellent reliability (.92). As was the case in the Reading test, the addition of five easy arithmetic items at the 8th grade increased the estimated reliabilities for Blacks over that of the originally proposed form. When the standard errors of measurement are compared there is virtually no difference between ethnic groups at any of the grade levels. There simply appears to be less heterogeneity among the Black examinees than among the Whites.

History and Government

Table 6 presents the reliabilities and group statistics for the revised single final form for the three grade levels. The reliabilities of the 30

****
Insert Table 6 about here
****

item revised form are good for a test of this length and show considerable improvement for Blacks at all grade levels over those of the originally proposed form. The shortened test form showed slightly larger differences between means across grade level in terms of standard deviation units than did the total item pool. This was particularly true in the transition between the 10th and 12th grade. The shortened form seems relatively sensitive to growth in that there is about 56% of a standard deviation gain between grades 8 and 10 and 63% of a standard deviation gain between grades 10 and 12.

Science

Table 7 repesents the reliabilities of this 25 item test and the mean contrasts with sex and ethnic groups.

****
Insert Table 7 about here
****

Inspection of Table 7 suggests that the 25 item test was equally as reliable as the original 42 item test. The estimated reliabilities are lower than the other three tests and are consistent with the finding of a multidimensional test.

Conclusions

The four revised 8th grade shortened forms, which are currently being administered in the field, included 21 reading items, 40 mathematics items, 30 social studies items, and 25 science items. All items in the pretest item pools were reviewed and approved by a sensitivity committee with respect to potential sex and/or ethnic bias. One item in the 8th grade Mathematics form was replaced because of potential content bias favoring boys. The Reading and Mathematics tests are projected to be "adaptive" according to grade level. That is, students in the 8th, 10th and 12th grade receive slightly different forms on each retesting. There is, however, sufficient item overlap across grades to make adjustments in IRT scaling parameters if needed. The change in the 10th and 12th grade Reading and Mathematics forms was proposed in order to minimize the impact of ceiling effects on gains. The use of adaptive tests in Reading and Mathematics assumes IRT scaling. Both the proposed Reading and Mathematics tests met the assumptions necessary for IRT scaling.

The History and Government and Science tests were less homogeneous in content, and in the case of the Science test, apparently not appropriate for IRT scaling. The History & Government test appeared to have little potential problems with either ceiling or floor effects, and it is suggested that one form be used at all grade levels. It is also suggested that one form of the apparently multidimensional Science test be used at all grade levels.

Reference

Rock, D.A. & Pollack, J.M. (1987) Measuring gains - A new look at an old problem. Paper given at the Multivariate Behavioral Research annual conference, November; Vancouver.

Table 4
SCORE STATISTICS BY GRADE, SEX AND RACE
READING TEST

	TOTAL	MALE	FEMALE	BLACK	WHITE	HISPANIC	ASIAN	EFFECT SIZE* MALE-FEMALE	WHITE-BLACK	WHITE-HISPANIC	WHITE-ASIAN
GRADE 8											
Number of Cases	700	321	318	92	375	126	84				
All Items:											
Mean, Effect Size	25.8	24.1	27.5	20.5	27.8	23.4	26.7	-0.3	0.7	0.4	0.1
S.D.	9.9	10.0	9.6	8.6	9.5	9.5	10.1				
Alpha	0.90	0.90	0.90	0.87	0.90	0.89	0.91				
Selected Item Subset:**											
Mean, Effect Size	12.5	11.6	13.4	10.0	13.5	11.3	12.8	-0.4	0.7	0.5	0.1
S.D.	4.7	4.8	4.6	4.5	4.5	4.5	4.9				
Alpha	0.83	0.83	0.83	0.80	0.82	0.80	0.85				
GRADE 10											
Number of Cases	470	250	212	38	316	71	23				
All Items:											
Mean, Effect Size	30.6	29.3	32.0	23.6	31.6	30.2	31.2	-0.3	0.8	0.1	0.0
S.D.	10.3	10.9	9.4	8.6	10.4	9.1	10.5				
Alpha	0.92	0.93	0.91	0.87	0.92	0.89	0.93				
Selected Item Subset:**											
Mean, Effect Size	12.5	12.1	13.0	9.4	13.0	12.1	13.0	-0.2	0.7	0.2	0.0
S.D.	4.8	5.1	4.4	4.3	4.8	4.3	4.4				
Alpha	0.85	0.87	0.82	0.79	0.86	0.81	0.83				
GRADE 12											
Number of Cases	429	188	236	63	266	62	23				
All Items:											
Mean, Effect Size	33.0	31.8	34.0	30.9	33.9	33.9	27.5	-0.2	0.3	0.0	0.6
S.D.	10.4	11.7	9.1	10.7	10.4	8.6	10.1				
Alpha	0.93	0.94	0.91	0.93	0.93	0.90	0.91				
Selected Item Subset:**											
Mean, Effect Size	12.8	12.4	13.0	11.7	13.1	13.1	11.2	-0.1	0.3	0.0	0.4
S.D.	4.7	5.1	4.4	4.9	4.7	4.0	5.0				
Alpha	0.85	0.88	0.83	0.86	0.86	0.80	0.87				

* EFFECT SIZE = (SUBGROUP1 MEAN - SUBGROUP2 MEAN)/TOTAL GRADE STANDARD DEVIATION

** SUBSET OF ITEMS SELECTED FOR TEST FORMS:

Grade 8: 1,2,3,4,5,16,17,18,20,21,22,24,26,28,30,31,32,33,35,36,37

Grade 10: 6,7,8,9,10,11,12,16,17,18,20,21,22,23,24,26,28,35,36,37

Grade 12: 6,7,8,9,16,17,18,35,36,37,39,40,41,42,45,46,47,48,49,50

Table 5
SCORE STATISTICS BY GRADE, SEX AND RACE
MATHEMATICS TEST

	TOTAL	MALE	FEMALE	BLACK	WHITE	HISPANIC	ASIAN	EFFECT SIZE* MALE-FEMALE	WHITE-BLACK	WHITE-HISPANIC	WHITE-ASIAN
GRADE 8											
NUMBER OF CASES	667	311	272	125	303	127	84				
ALL ITEMS:											
Mean, Effect Size	30.7	31.4	29.5	24.4	33.9	28.2	33.9	0.2	0.7	0.5	0.0
S.D.	12.7	13.2	12.0	8.9	13.4	11.0	13.5				
Alpha	0.90	0.91	0.89	0.80	0.91	0.87	0.92				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	19.5	19.9	18.8	15.0	21.5	18.0	22.2	0.1	0.7	0.4	-0.1
S.D.	8.7	9.0	8.4	6.8	9.0	7.8	8.8				
Alpha	0.90	0.91	0.89	0.83	0.91	0.87	0.90				
GRADE 10											
NUMBER OF CASES	599	225	368	88	346	88	34				
ALL ITEMS:											
Mean, Effect Size	42.1	45.2	40.5	28.4	46.7	37.5	51.2	0.3	1.2	0.6	-0.3
S.D.	15.7	16.5	14.9	12.2	14.3	14.0	18.1				
Alpha	0.94	0.95	0.93	0.90	0.93	0.92	0.96				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	25.6	27.0	25.0	17.0	28.5	23.0	29.9	0.2	1.3	0.6	-0.1
S.D.	9.2	9.0	9.0	7.6	8.1	8.7	8.6				
Alpha	0.92	0.92	0.92	0.86	0.91	0.90	0.93				
GRADE 12											
NUMBER OF CASES	546	210	336	69	351	73	36				
ALL ITEMS:											
Mean, Effect Size	49.1	52.6	46.9	32.9	53.4	41.5	54.2	0.4	1.3	0.7	0.0
S.D.	16.2	16.1	15.9	12.7	14.6	15.0	14.9				
Alpha	0.95	0.95	0.95	0.90	0.94	0.93	0.95				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	22.8	24.5	21.7	14.2	25.0	18.7	25.6	0.3	1.2	0.7	-0.1
S.D.	9.1	9.0	9.0	6.8	8.5	8.1	8.1				
Alpha	0.92	0.92	0.91	0.83	0.91	0.88	0.91				

* EFFECT SIZE = (SUBGROUP1 MEAN - SUBGROUP2 MEAN)/TOTAL GRADE STANDARD DEVIATION
** SUBSET OF ITEMS SELECTED FOR TEST FORMS:
Grade 8: 1,2,4,5,7,8,9,10,14,15,16,17,18,22,25,28,29,32,35,36,38,40,41,42,43,44,45,46,47,48,51,53,60,68,73+3 HSB+2 NEW
Grade 10: 1,2,4,5,7,8,9,10,12,14,15,16,17,18,22,25,27,28,29,31,32,34,35,36,38,40,41,42,43,44,45,46,47,48,51,53,54,60,68,73
Grade 12: 1,4,6,7,8,9,13,15,16,17,18,19,20,22,34,38,40,43,44,49,50,51,55,56,58,60,61,64,65,66,69,71,72,74,75,78,80,81+2 HSB

Table 6
SCORE STATISTICS BY GRADE, SEX AND RACE
History and Government

	TOTAL	MALE	FEMALE	BLACK	WHITE	HISPANIC	ASIAN	EFFECT SIZE* MALE-FEMALE	WHITE-BLACK	WHITE-HISPANIC	WHITE-ASIAN
GRADE 8											
NUMBER OF CASES	669	311	273	123	306	127	85				
ALL ITEMS:											
Mean, Effect Size	31.7	31.4	31.7	26.4	33.4	31.5	33.9	0.0	0.7	0.2	0.0
S.D.	10.0	10.3	9.7	8.7	10.0	9.6	9.4				
Alpha	0.89	0.89	0.88	0.84	0.89	0.87	0.88				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	17.1	16.9	17.1	14.0	18.2	16.8	18.4	0.0	0.7	0.2	0.0
S.D.	6.1	6.2	5.9	5.5	6.2	5.8	5.5				
Alpha	0.86	0.87	0.85	0.81	0.87	0.84	0.83				
GRADE 10											
NUMBER OF CASES	769	289	470	110	427	140	40				
ALL ITEMS:											
Mean, Effect Size	36.4	37.3	35.8	27.7	38.6	36.0	38.3	0.2	1.1	0.3	0.0
S.D.	9.5	10.7	8.5	7.7	9.2	8.0	9.4				
Alpha	0.89	0.92	0.86	0.80	0.89	0.84	0.89				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	20.4	20.5	20.3	15.1	21.7	20.3	21.4	0.0	1.2	0.2	0.0
S.D.	5.6	6.4	5.1	5.1	5.3	4.9	5.4				
Alpha	0.86	0.90	0.83	0.79	0.86	0.81	0.86				
GRADE 12											
NUMBER OF CASES	699	255	442	81	430	123	42				
ALL ITEMS:											
Mean, Effect Size	42.7	43.7	42.2	30.9	44.9	42.4	44.1	0.1	1.4	0.2	0.1
S.D.	10.2	10.9	9.8	10.7	8.8	9.2	9.2				
Alpha	0.92	0.93	0.91	0.91	0.90	0.89	0.90				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	23.9	24.1	23.8	17.7	25.1	23.8	24.4	0.1	1.3	0.2	0.1
S.D.	5.6	6.0	5.4	6.5	4.7	5.1	5.2				
Alpha	0.90	0.92	0.88	0.90	0.87	0.86	0.89				
MEAN CHANGE (POOLED S.D. UNITS)											
Grade 8 to Grade 10	0.56	0.57	0.60	0.21	0.61	0.64	0.56				
Grade 10 to Grade 12	0.63	0.58	0.66	0.45	0.67	0.70	0.56				

* EFFECT SIZE = (SUBGROUP1 MEAN - SUBGROUP2 MEAN)/TOTAL GRADE STANDARD DEVIATION

** SUBSET OF ITEMS SELECTED FOR TEST FORMS:
All Grades: 2,3,6,7,9,12,15,16,17A,17B,21,23,24,25,26,29,30,32,33A,33B,33C,33E,33G,36,38,39,40,42,43,47

Table 7
SCORE STATISTICS BY GRADE, SEX AND RACE
Science Test

	TOTAL	MALE	FEMALE	BLACK	WHITE	HISPANIC	ASIAN	EFFECT SIZE* MALE-FEMALE	WHITE-BLACK	WHITE-HISPANIC	WHITE-ASIAN
GRADE 8											
NUMBER OF CASES	697	319	317	91	373	126	84				
ALL ITEMS:											
Mean, Effect Size	16.6	17.1	16.2	13.9	17.7	14.8	17.4	0.2	0.7	0.6	0.1
S.D.	5.3	5.7	4.8	4.1	5.2	4.7	5.4				
Alpha	0.69	0.74	0.64	0.50	0.68	0.62	0.71				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	10.4	10.8	10.1	8.1	11.2	8.9	11.3	0.2	0.8	0.6	0.0
S.D.	4.1	4.4	3.9	3.3	4.1	3.6	4.3				
Alpha	0.70	0.74	0.66	0.55	0.68	0.59	0.73				
GRADE 10											
NUMBER OF CASES	473	253	212	39	316	72	24				
ALL ITEMS:											
Mean, Effect Size	20.1	21.3	18.6	16.7	21.2	16.8	21.3	0.4	0.7	0.7	0.0
S.D.	6.5	6.7	5.9	6.5	6.5	4.6	6.7				
Alpha	0.80	0.82	0.76	0.81	0.80	0.60	0.83				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	13.2	14.1	12.0	10.4	14.1	10.6	13.8	0.4	0.7	0.7	0.1
S.D.	5.0	5.1	4.6	5.4	4.9	3.6	5.3				
Alpha	0.80	0.82	0.76	0.85	0.80	0.60	0.84				
GRADE 12											
NUMBER OF CASES	424	187	233	62	265	61	23				
ALL ITEMS:											
Mean, Effect Size	21.2	22.7	20.1	16.9	22.7	19.8	19.7	0.4	0.9	0.4	0.5
S.D.	6.6	7.2	5.8	5.1	6.5	5.1	7.9				
Alpha	0.81	0.85	0.76	0.67	0.81	0.68	0.88				
SELECTED ITEM SUBSET:**											
Mean, Effect Size	14.2	15.4	13.4	11.1	15.5	13.0	13.0	0.4	0.9	0.5	0.5
S.D.	4.8	5.0	4.5	4.1	4.6	4.2	5.7				
Alpha	0.79	0.82	0.75	0.68	0.78	0.72	0.86				
MEAN CHANGE (POOLED S.D. UNITS)											
Grade 8 to Grade 10	0.61	0.70	0.45	0.57	0.63	0.47	0.53				
Grade 10 to Grade 12	0.22	0.25	0.32	0.16	0.30	0.60	-0.14				

* EFFECT SIZE = (SUBGROUP1 MEAN - SUBGROUP2 MEAN)/TOTAL GRADE STANDARD DEVIATION

** SUBSET OF ITEMS SELECTED FOR TEST FORMS:
All Grades: 3,4,5,6,9,10,11,12,14,16,17,18,20,21,22,23,24,26,28,29,30,33,36,41,42

Draft

METHODOLOGICAL EXPERIMENTS OF THE NELS:88 FIELD TEST STUDENT SURVEY

Samuel R. Lucas

NORC / University of Chicago

This paper is being presented at the AERA Annual Meeting in New Orleans on April 8, 1988. The views expressed in this paper are not necessarily those of either NORC or the Center for Education Statistics of the U.S. Department of Education.

I. Introduction: Choosing the "Best" Respondent

One advantage of the multi-respondent design used in NELS:88 is that particular information may be asked of only the best respondent population, if this population can be identified. The designation of a best respondent population for a given question is not as straightforward as one might imagine. Indeed, given the difficulty in designating one "best" respondent, one might ask whether or not it would be more advantageous to simply ask all respondents the same question. Some questions, attitudinal questions, or questions which will reveal important systematic variation in the apprehension of phenomena of key importance to the research may usefully be asked of more than one respondent. Still, the underlying assumption for data collection, adopted because of various constraints (typically respondent burden and available administration time) is that a minimum number of duplicate questions will be asked. For basic factual questions, such as SES, the student's educational history, and similar questions, the identification of and reliance upon one best respondent will benefit the research by freeing other respondents for more questions on the salient topics of the research.[1] Thus, where possible, it is valuable to pursue the designation of one "best" respondent.

Several factors influence the designation of a particular respondent as "the best" respondent for a given item. With a study design such as that employed by NELS:88, a few of the factors considered were:

---

[1]It is worth noting that for some research even these baseline questions may usefully be asked of more than one respondent. For example, a study to examine the assessment of SES by various actors might ask four respondents with different relationships to an individual the same question. Thus, as in most cases, the research objectives drive the research design.

1. Proportion of respondents surveyed to respondents for which data will be needed.
2. Respondent's access to the relevant information.
3. The amount of time available for survey administration.
4. Whether or not designation of the item as critical will be advantageous.

On NELS:88 the first concern was exemplified in the difference between the number of parents surveyed compared to the number of students surveyed. For example, if the study design had called for 26,000 students to be sampled, but perhaps involved a sub-sample of only 15,000 parents, questions involving key data such as SES could not be confined to the parent questionnaire. However, in NELS:88, parents of each sampled student will be asked to respond to the parent questionnaire in the base year.[2]

The second criterion, the respondent's access to the relevant information is also of considerable importance. Students are far less likely than parents to know their parents income and financial investments, if any. Still, the researcher may consider various steps to insure that, on the day of survey administration, the respondent will have the information. For example, at one time thought was given to requesting that parents tell their children the family income on the day prior to survey day. For various reasons, this plan was rejected; however, its discussion illustrates some of the options available to counter the limited knowledge of a respondent.

---

[2]The study will not achieve a one-to-one correlation between parent and student because parents of sampled students in state augmentation schools will not be asked to respond to the parent questionnaire. In addition, the parent study may achieve lower response rates than the student. These issues complicate the criterion of the number of various respondents surveyed, and must be considered for a variable of such great importance as SES.

For group administered, telephone administered, and, perhaps to a lesser extent, personal (face-to-face) administered surveys the issue of time required to complete the document is a significant consideration. For a study in institutions, where interaction with gatekeepers entails a priori limitation and standardization of the time available across respondents despite the variation in the actual time required to administer the survey, the importance of devising a document which can be completed by those least able to complete the instrument is of crucial significance for the collection of unbiased data. Where the ability of the respondent to complete the instrument is systematically related to the issue under study, as in an education study, the time available takes on even greater importance. In this situation, moving questions from the instrument targeted to the respondent with the greatest time limitation to other instruments can serve to increase the value of the data across the board.

Lastly, designation of an item as critical, which will cause the item to be retrieved for cases where it is missing, may also be considered in lieu of removing a question from a particular instrument. Designation of a question as critical for retrieval should decrease the non-response rate, if the respondent actually has the information requested.

Deciding to designate a "best" respondent, therefore, will involve weighing the sampling plan, the types of respondents and their access to the required information, the amount of time available for survey administration, and the options afforded by designating some items critical for retrieval. After considering these issues, however, the job is only partially done. The designation of a particular respondent as the best one for a given question is itself a complicated question. If possible, efforts should be made to avoid the arbitrary selection of a "best" respondent or at least to check the designation

by analyzing the data obtained from various sources. Of course, in the final analysis, some source of data must be used as the standard by which others are measured; the selection and validation of this standard, where possible, can greatly enhance the value of the ultimate data collection strategy adopted.

Two decisions to adjust the NELS:88 student instrument provide a clear illustration of the ways in which the considerations detailed above may interact. The first decision concerned the development of a strategy to successfully obtain SES data, the second decision concerned the identification of a best respondent for information concerning the physical and mental disabilities of students, and the third decision concerned the identification of the best respondent to answer questions concerning the early educational experiences of the student. Each of these decisions are discussed below.

II. Socioeconomic Status

For the Field Test, we obtained information on parental occupation and education from both parents and students. In addition, we obtained student responses to an item list and to a question concerning the number of rooms in the household. We also obtained parental income information--both salary and investment/business income--in the parent questionnaire. The questions and respondents that might be used to derive a measure of SES are listed below in table S-1:

TABLE S-1

POSSIBLE SES COMPOSITE VARIABLES

Question	Student Respondent	Parent Respondent
Close-ended Parents Occupation	Yes	Yes
Open-ended Parents Occupation	Yes	Yes
Parents Education	Yes	Yes
Item List	Yes	No
Rooms in Household	Yes	No
Parent Income	No	Yes

We analyzed the data in several ways. We drew a subsample of cases and compared responses to all formats of the parent occupation question, to determine whether or not parents and students were effectively utilizing the close-coded categories. We also conducted an analysis of nonresponse, to determine whether or not particular questions were causing a large number of problems. Lastly, we paired up student and parent data to determine the correlation between the parent and student derived reports of the SES composite.

Each portion of our analysis--the comparison of open-ended and close-coded responses, the analysis of nonresponse, and the comparison of the correlation of parent and student SES measures--utilizes a slightly different criterion of effectiveness. It was important to employ varying definitions of effectiveness because any other method would have relied upon the arbitrary selection of a particular respondent as "best" and comparisons of other respondents to the arbitrarily selected one. We did not want to simply assume that the parent would be a better respondent than the student when it came to SES data.

The first part of our analysis, the comparison of open-ended and close-ended responses, aimed at certifying the parent as a valid respondent, deciphering the cause of any student difficulty discovered, and providing a preliminary comparison of student and parent responses which might inform later comparisons. Our procedures in this part of our analysis were quite straightforward. We had a coder steeped in the issues with which the study was concerned and the various perspectives on SES code the open-ended data in the student and parent questionnaires. We then compared the professional coding to the respondent coding. We analyzed 550 cases in this manner.[3]

The results of this analysis were threefold. The analysis suggests that the parent was indeed a better respondent than the student, that the parent was able to effectively utilize the close-coded categories, and that the student was able to effectively utilize an open-ended approach to obtaining parent occupation information.

The parent was a better respondent than the student when the amount of agreement between the professional coder, utilizing the respondents open-ended statement, and the respondent's coding are compared. It should be remarked that the comparisons are based on the information provided by the respondent to which the comparison will be made. For example, a comparison is made between the student's closed-coded response and the professional coding derived from the student's open-ended response. In this example, the open-ended information from the parent does not inform the coder's classification. This approach was adopted because the goal of the exercise was to determine whether or not the respondent was able to translate the information in their possession into the close-coded

---

[3]This analysis was conducted by practicum student Tim Ryan under the supervision of Steven Ingels; results herein mentioned are drawn from Ryan's paper.

categories effectively. The approach adopted insured that the coder would make a judgement with information similar to that of the student, and information similar in type to that which we might receive in the base year. The same approach was used when it came to comparing parent respondent coding with professional coding. The results of each of the comparisons are below, in table S-2:

TABLE S-2

COMPARISON OF ASSIGNED CODING

Father

	Student			Parent		
	#	Raw%	Valid%	#	Raw%	Valid%
Same codes	258	47.1	66.2	255	54.0	72.4
Different codes	132	24.1	33.8	97	20.6	27.6
Not counted	158	28.8		120	25.4	
	548			472		

Mother

	Student			Parent		
	#	Raw%	Valid%	#	Raw%	Valid%
Same codes	277	50.4	72.1	278	58.9	73.9
Different codes	107	19.4	27.9	98	20.8	26.1
Not counted	166	30.2		96	20.3	
	550			472		

The data above suggest that the parent was better able to utilize the close-coded categories than the student. In addition, the amount of agreement between the coder and the parent is high. Lastly, the amount of agreement between the student and the coder is not as low as one might expect.

A check of the amount of agreement between parent data, coded by the respondent, and student data, coded by a professional, revealed a marginal improvement in the student data. In addition, the coder reported that oftentimes the student isolated a particular feature of the parents' occupations and coded the occupation accordingly. For example, one student noted that their mother worked as a cashier and their father was a cook, when in fact the parents owned the business. This qualitative data was added to the growing body of quantitative data and informed the next two stages of our analysis and the process of revising the instruments.

The analysis of nonresponse aimed at determining whether or not the difficulties in devising an SES measure were more related to the particular questions and the wording used than to the underlying construct being measured. The data were obtained in different forms in the parent and student questionnaires. Still, the data reported in table S-3 below provide some indication of the differing levels of nonresponse for each of the respondents on similar questions:

TABLE S-3

NONRESPONSE RATES OF PARENT AND STUDENT REPORTS

OF PARENTAL OCCUPATION AND EDUCATION

Question-Respondent	DK	MR	Blank	Ref	Total Nonresponse
Father's occupation-Student	8.4%	8.6%	7.6%	.3%	24.9%
Mother's occupation-Student	6.0%	7.5%	5.2%	.2%	18.9%
Respondent's occupation-Parent	3.0%	4.4%	5.9%	.4%	13.7%
Spouse's occupation-Parent	.7%	2.6%	20.3%	---	23.6%
Father's education-Student	21.4%	6.7%	6.7%	---	34.8%
Mother's education-Student	18.5%	7.2%	4.8%	---	30.5%
Respondent's education-Parent	---	10.3%	.5%	.3%	11.1%
Spouse's education-Parent	---	11.2%	5.4%	---	16.6%

Comparatively high rates of nonresponse are recorded for each of the items in the student questionnaire. The data suggests that students are severely handicapped in their ability to provide the requested information. Though there is some variation that might be remedied by revisions of the question wording, the large incidence of student DK responses, especially on the parental education questions, seems to indicate that changes in the wording of the question will have only a slight effect on the ability of students to respond.

The last part of our analysis called for treating the occupation and education data in a manner similar to its use by many analysts, namely, in a composite variable designed to measure SES. Without entering the involved debate concerning the particular concepts which underlie the construct, we devised our composites using only the data which were identical across all survey

instruments. Therefore, we only included the parental occupation and education data. We were able to use other information also obtained--responses to an item list, the number of rooms in the household, and family income--as a check on the SES measure we devised.

Given that earlier analyses had tended toward affirming our intuitive assessment that the parent would be the better reporter of SES data, we treated the parent derived measure of SES as a yardstick with which to measure the amount of student deviation. The correlation matrix for the two SES measures and related variables, with data derived from the parent questionnaire in bold, is below, in table S-4[4]:

TABLE S-4

CORRELATION MATRIX OF PARENT AND STUDENT DERIVED SES MEASURES

	Stu-SES	**Salary**	**Busi-farm**	Rooms
Par-SES	.7361**	.4805**	.2573**	.3235**
Stu-SES		.3875**	.1580	.2718**
Salary			.4460**	.3543**
Busi-farm				.3000**

**=.001

The composite SES measure derived from the parent data correlates more highly with other measures of SES, such as family income (both salary and business/investment income), and number of rooms in the house. The data seemed to suggest that utilization of a measure of SES derived from student reported

[4]The individual items which composed the item list were themselves tested during this juncture of the analysis. Correlation data for each of the individual items has been omitted.

data education and occupation data should be an option of last resort.

The analysis of field test data suggested that we should alter our strategy for obtaining student reports of SES in some fashion. Indeed, the data seemed to suggest that we should consider removing the parental occupation and education questions entirely from the student instrument. Still, other considerations discussed above had to be weighed in coming to a final decision. Our three types of analyses led us to draw the conclusions listed below, in Table S-5:

TABLE S-5

RESULTS OF ANALYSIS OF STUDENT AND PARENT REPORTS ON SES VARIABLES

Source	Conclusion
Professional Coding	Parents effectively used the close-coded categories
	Students were able to effectively respond with the open-ended questions
Nonresponse analysis	Parent nonresponse rates were lower than student nonresponse rates
	Alterations in the parental education question would marginally affect student nonresponse rates
Composite Comparisons	Parent-derived SES measures were more highly correlated with other measures of SES than the student-derived measure

This set of information pointed toward greater reliance on the parent for occupation and education data. Still, that students were able to effectively respond to open-ended questions on parental occupation argued for the retention of an open-ended parental occupation question in the student questionnaire. Retention of the open-ended question would allow analysts to use either the parent derived data or, in the case of missing parent data, rely on the best means of obtaining parental occupation from students.

Given that the inclusion of several state supplements will mean a substantial number of student cases will not have accompanying parent data, and owing to the centrality of SES for a number of analyses, this approach seemed best. Thus, our effort to identify a best respondent in this case led to a mixed approach. Still, our efforts were not in vain. The amount of respondent confusion on the part of students attempting to use the close-coded categories, the amount of retrieval time saved should substantially increase the quality of other data, since survey personnel will be able to focus on data elements more amenable to the retrieval effort.[5]

III. Student Physical and Mental Disabilities

For the Field Test we obtained reports on the student's physical and mental disabilities from both parents and students. In addition to obtaining the information from two different respondents, we also used the opportunity to test the value of various question wordings by asking each respondent the question in two different ways. The different forms of the question employed, however, phrased the question with different time orientations. Still, comparisons were possible across all versions of the questions employed.

Main Questionnaire Question 9 asked students "Do you or have you received special services for any or all of the following conditions?", and conditions

---

[5]Another NELS:88 experiment called for in-house quality control checks of on-site retrieval efforts. One result of this additional experiment was the generation of retrieval rates on a per-question basis, and rating of retrieval errors. Twenty questions had been designated critical for the field test, among them the close-coded parental occupation question. The parental occupation question accounted for 11.4% of the retrieval burden, the second highest retrieval burden imposed. Deletion of this question, therefore, will greatly increase the amount of time available for retrieval on other items.

were listed and space provided for a yes or no response. Main Questionnaire Question 48 asked students "Are you enrolled in any of the following special programs/services?", and conditions were listed below with space for a yes or no response. The format of the questions used in the parent questionnaire was virtually the same, and introduced the same type of now/ever dichotomy. Tables S-6 and S-7 below list the two special services/conditions which were compared:

TABLE S-6

ORTHOPEDIC IMPAIRMENTS

Respondent	Time Frame	Text
Student	Ever	Orthopedically impaired (for example, club foot, absence of arm or leg, cerebral palsy, amputation, polio)
Student	Now	Special services for the orthopedically handicapped
Parent	Ever	Orthopedic problem (for example, club foot, absence of arm or leg, cerebral palsy, amputation, polio)
Parent	Now	Special services for orthopedically handicapped

TABLE S-7

LEARNING DISABILITIES

Respondent	Time Frame	Text
Student	Ever	Specific learning disability (for example, imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations)
Student	Now	Special Education services for students with learning problems
Parent	Ever	Specific learning problem (for example, imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations)
Parent	Now	Special education services for slow learners

In the case of both orthopedic problems and learning disabilities we crosstabulated responses to assess whether or not respondents had contradicted themselves in answering the two different versions of the question. Next, we compared student and parent responses. Lastly, in the case of orthopedic problems, we re-contacted a subsample of suspicious cases and re-asked the question. Again, our methods reflected our desire to avoid the arbitrary selection of one respondent as the best one without at least attempting to empirically verify our intuitions.

We shall limit our discussion to the issue of orthopedic problems, since our findings in both cases were quite similar. The comparison of student responses to the questions on orthopedic problems resulted in the crosstabulation below, in table S-8:

TABLE S-8

CROSSTABULATION OF STUDENT RESPONSES

	Now		
Ever	Yes	No	Total
Yes	1	33	34
No	30	1179	1209
Total	31	1212	1243

The data in table S-8 shows that only 63 students out of the 1243 reported contradictory information, while 1179 students consistently marked "No" and only one student consistently marked "Yes". But given that only one student consistently responded "Yes", the 63 Yes/No responses really lead to the conclusion that, on the basis of the cross-validation of the student responses, up to 98.4% of the "Yes" responses were erroneous. The level of erroneous "Yes"

responses may have been as low as 51.6% since the thirty-three students who responded "No" to the question with a "now" orientation might have correctly recorded their departure from such programs in this manner. Still, the amount of "noise" in the data on the policy relevant group--those students receiving special services or having received special services--might be anywhere from 51.6% to 98.4%. The data received from parent reports presented a much more focused and therefore presumably more accurate picture, as table S-9 reveals:

TABLE S-9

CROSSTABULATION OF PARENT RESPONSES

	Now		
Ever	Yes	No	Total
Yes	0	8	8
No	0	703	703
Total	0	711	711

Ninety-nine percent of the parents indicated that their children were not now nor ever had been enrolled in programs for the orthopedically handicapped. The eight other parents may or may not have contradicted themselves. Technically, since all eight of the parents who responded affirmatively for their child's ever having received services for orthopedic handicaps also responded negatively, it is possible that 100% of the parents were in error. This was not regarded as likely, however, given the pattern of student responses. Still, direct comparisons of student and parent reports would allow us to address the issue.

Comparisons of student and parent responses on the strictly parallel items were conducted. The crosstabulation of responses on the question utilizing an ever orientation appears below in tables S-10. Most parents and students agreed. Still, a troublesome amount of disagreement existed, casting doubt upon the smattering of "Yes" responses:

TABLE S-10

CROSSTABULATION OF PARENT AND STUDENT RESPONSES

WITH AN "EVER" ORIENTATION

	Parent Report		
Student Report	Yes	No	Total
Yes	2	19	21
No	6	643	649
Total	8	662	670

Similar data were obtained for the question which asked whether or not the student now received services for an orthopedic problem. The amount of parent and student disagreement remained slight, but parents never corroborated their student's "Yes" responses. Table S-11 below is the crosstabulation of the "now" orientation question:

TABLE S-11

CROSSTABULATION OF PARENT AND STUDENT RESPONSES

WITH A "NOW" ORIENTATION

	Parent Report		
Student Report	Yes	No	Total
Yes	0	16	16
No	0	647	647
Total	0	663	663

In the light of this matrix of findings, we re-contacted a subsample of ten respondents whose answers on the ever orientation question were contradictory. The information garnered in these contacts indicated that parents and students interpreted the questions differently, as remarked in the Field Test Report:

> In most cases, the affirmative student reply . . . was based on a broken limb--a condition that the parent was equally aware of, but adjudged not chronic, and to be outside the intentions of the question. Here student error was abetted by the examples of orthopedic problems--club foot, absence of limbs, and so on--so that the student classed fractured or broken limbs as orthopedic impairments.[6]

On the basis of the analysis of field test data we provisionally designated the parent as the best respondent for items on student physical and mental disabilities. We also proposed alterations in the student version of the question to assist the student, such as including an example of a non-orthopedic problem. In addition, we raised the issue of whether or not school records might serve as an additional check on both parent and student responses.

We recommended, however, the deletion of the question from the student questionnaire. The quality of student data was so low as to make sole reliance upon the student report a dubious proposition, even in situations where no parent questionnaire was present in the datafile. Thus, though the data is of appreciable importance to policy-makers, the inability of eighth grade students to effectively respond to the questions (and the questionable ability of survey personnel to make any meqaningful changes in the question so as to remove the

---

[6]Ingels et al., Field Test Report: National Education Longitudinal Study of 1988 (NELS:88), (Chicago: NORC, 1987), p. 118.

difficulties) makes collection of the data from the student a waste of valuable survey administration time.

IV. Student early educational experiences

The analysis to determine the best respondent for student early educational activities was more straightforward than either of the two preceeding analyses, though no less important. For the Field Test we obtained reports of student early educational experiences from both students and parents. One five part question in each of the instruments asked whether or not the student had attended kindergarten, day care, nursery school or preschool, head start, or extended day. The levels of nonresponse recorded for each of the questions is provided below in table S-12:

TABLE S-12

NONRESPONSE RATES FOR EARLY EDUCATIONAL EXPERIENCES

	Parent			Student		
	DK	Blank	Tot NResp	DK	Blank	Tot NResp
A. Kindergarten	.4%	5.5%	5.9%	.9%	3.7%	4.6%
B. Day Care	1.0%	15.0%	16.0%	10.8%	12.2%	23.0%
C. Nursery school or preschool	.1%	10.1%	10.2%	8.1%	8.5%	16.6%
D. Head Start	1.9%	21.2%	23.1%	19.5%	14.0%	33.5%
E. Extended Day	3.2%	20.9%	24.1%	24.0%	15.0%	39.0%

For every question except kindergarten, the highest nonresponse rate was registered in the student questionnaire. Moreover, the sources of nonresponse were quite different, in the student case being primarily don't know responses and in the parent case being blanks. This pattern of responses suggested that the level of lost data in the case of the student would not be appreciably altered by changes in the research instrument. A comparison of the student and parent answers for was quite revealing. In the Field Test Report we remarked:

> . . . there is a fair amount of disagreement between parents and students on these items. Perhaps even more striking is the high degree of student uncertainty about some of the items, and the capacity of parents to answer those items which the student has responded Don't Know. Overall, parents used the Don't Know option 36 times. Students in this analysis (that is, those whose parent had also participated) used Don't Know 395 times, over ten times more often. Parents were generally able to give answers where students had opted for Don't Know--in 56B [Day care], for example, all 71 of the student Don't Knows have been answered yes or no by parents; for 56C [Nursery school or preschool], all 51 student Don't Knows are answered yes or no by parents; in 56D [Head Start], 115 of 118 student Don't Knows were resolved by parents, and in 56E, [Extended Day], 140 of 145.[7]

On the basis of this information we opted for removing the questions concerning early educational experiences from the student questionnaire. The need to shorten the field test instrument for use in the base year coupled with the better quality of the parent data on these questions led to the decision.

---

[7]Ibid., p. 155.

Even though some parents may not be asked to complete a questionnaire, we felt that to attempt to obtain data on early educational experiences from all 26,200 students would pose an undue burden upon the vast majority of student respondents.

V. Conclusion

The experiments discussed above served as examples of the effort to identify a particular "best" respondent for specific questions. These efforts extended to all four respondent populations--students, parents, teachers, and principals. Many other experiments of the same type were conducted, involving the report of ethnicity information, locating data (particularly involving the expected tenth grade school), and more.

The type of experimental approach adopted during the NELS:88 Field Test, though initially costly, results in a lower final response burden, higher quality data, and, by allowing the re-assessment of intuitively apt preliminary decisions, possibly even the advance of particular aspects of social science research.

METHODOLOGICAL EXPERIMENTS OF THE NELS:88 FIELD TEST PARENT SURVEY

Betty Bower Simon NORC / University of Chicago

NELS:88 Parent Survey Field Test results show that when parent questionnaires are delivered to parents by their eighth graders as compared to U.S. mail: the percentage of remailing required is similar; pre-prompt and final questionnaire completion rates are higher; and the quality and quantity of response to questions soliciting personal locator and income information is superior. Not predesignating a particular parent as respondent, but rather instructing that the questionnaire be completed by the parent most knowledgeable about the eighth grader's education, resulted in about 75 percent of the questionnaires being completed by mothers, about 19 percent being completed by fathers. Significant differences were found in the way in which male parents and female parents answered several attitudinal and behavioral questions. Furthermore, approximately 33 percent of the eighth graders were found not to be living with both of their biological parents. Investigation confirmed considerable confusion of reference in the cases of single-parent households or remarried-parent households. When questioned about a father (male guardian) or mother (female guardian) in the student survey, eighth graders often described an out-of-house parent, sometimes even when a step-parent lived in the child's household. The implications of the parent and the child potentially describing respectively a step-parent and the biological parent are discussed with reference to creating education-relevant measures of socio-economic status. It is suggested that one question in particular — whether longitudinal studies purporting to study parental influence on educational development should survey two (or more) parents of each sampled child — merits further investigation.

Two principal objectives of the Field Test of the NELS:88 Parent Survey were: (1) to determine whether Parent Questionnaires were best sent to the homes by eighth-grader delivery or by the U.S. Mail; and (2) to examine the effects of not pre-designating a particular parent as the study's respondent (the parent questionnaire packet was addressed: TO THE PARENTS OF (NAME OF CHILD) -- rather than to a specific parent, instructions being provided in an accompanying cover letter that "the questionnaire should be completed by the parent or guardian who is most familiar with the student's current school situation and educational plans.") Basically, we wanted to know: Will eighth grade students leave the Parent Questionnaires they have been given to take home at the bottom of their lockers? Will only mothers fill out the questionnaires anyway or is there some tangible payoff to giving the father the chance to be considered as the parent who is more knowledgeable about the eighth grader's education? While investigating these matters, we stumbled into a minefield of intriguing issues, the sum of which can roughly be encompassed by the question: Who are the "parents" of these NELS:88 eighth graders after all?

While we will begin by summarizing the fairly unexpected results of the questionnaire distribution experiment, mainly our attention will be focused on the matter of identifying the NELS:88 parents. In so doing, we will address the following: What are the implications of not pre-designating a particular parent respondent? When an eighth grader has parents living in two different households, will s/he dependably deliver the questionnaire to the household of his/her principal residency? If the child has a parent living outside of the child's household, particularly if the custodial parent has remarried, when the custodial parent and the child are describing the child's second "parent", will both be describing the same individual? That is, will the child tell you all about his/her other biological parent, while the parent tells you about the step-parent? How are measures of socio-economic status complicated by such potential confounding, and is this necessarily a bad outcome? What are the implications of surveying only one of the child's parents; that is, since the mother and father often assume different roles in the family, should we be concerned that the child's mother and father may have different perceptions and attitudes with regard to their child's education and the child's future? Is there apparent justification for surveying two (or in some cases even more) of the child's parents in a future follow-up Parent Survey? Finally, given the high rate of divorce and remarriage, and given the possibility that even in intact families, the other parent may choose to be the respondent in a follow-up study, what are we losing by not systematically measuring a particular parent's change in behavior and attitude over time?

Below we will provide descriptive statistics and crosstabulations of the NELS:88 Field Test Parent Survey to argue that given the social reality -- many children do not live all of their childhood with the two biological parents -- and given substantial differences in the way that male and female parents do tend to answer particular behavioral, perceptual, and attitudinal questions of the NELS Parent Questionnaire, serious consideration should be given to surveying multiple parents of each NELS:88 student respondent.

The First Experiment: Optimal Method of Delivering Parent Questionnaire

Thirty-four of the 51 NELS:88 Field Test eighth-grade schools were selected to participate in the Parent Survey. To assure adequate representation of private schools, all 14 of the sampled private eighth-grade schools were designated to have the parent study. Schools located in areas with heavy student survey activity comprised the balance of the parent sample institutions. One parent/guardian of each of the selected eighth-grade students in the 34 schools was made a respondent in the study.

All parents received the Parent Questionnaire in a packet that also contained a letter introducing the study to the parent and a postage-paid envelope for return of the questionnaire to NORC. Three

different methods were used to deliver the questionnaires to parents: the eighth grader carried the packet home (Student Take Home); the packet was mailed from the school (School Mail); or the packet was mailed from NORC in Chicago (Central Office Mail). An important goal of the field test was to test the relative effectiveness of the Take Home and Mail treatments according to the following criteria:

- degree of remailing required
- short-term and longer-term parent response rates
- quantity and quality of response to selected critical items.

Student Take Home--the method that had students carrying the questionnaires home to their parents on the day that the students themselves were surveyed--emerged as a superior method for both maximizing parent participation and assuring high quality responses.

Criterion 1: degree of mailing required. Regarding the first of the three criteria, remailings made, Student Take Home showed 34 of 397 (8.6 percent) remailings, while School Mail showed 15 of 156 (9.6 percent) and Central Office Mail showed 26 of 244 (10.7 percent). This evaluation encompasses a subset of total parent cases, specifically those in which students were surveyed and no particular complications in parent questionnaire delivery were encountered. (Note: These results have experimental validity only for the Student Take Home and Central Office Mail treatments since assignment to the School Mail treatment was by self-selection. School Mail was used in only five schools. It was used when the school refused to provide parent address information that was necessary for the Central Office Mail treatment.)

REQUIRED REMAILINGS UNDER THE THREE DELIVERY METHODS

	number	percentage
Student Take Home	34/397	8.6%
School Mail	15/156	9.6%
Central Office Mail	26/244	10.7%

Criterion 2: short-term and long-term response rates. With regard to the second criterion, early and late participation rates, Student Take Home consistently emerged as more effective than Central Office Mail.

At three weeks, the time at which telephone prompting of outstanding questionnaires was scheduled to begin, 53.3 percent of the Student Take Home questionnaires and 47.7 percent of the Central Office Mail questionnaires were in house. (This analysis encompasses the same set of cases examined in the remailing analysis.)

PARENT PARTICIPATION RATES (PRE-PROMPT) WITH STUDENT TAKE HOME AND CENTRAL OFFICE MAIL

	number	percentage
Student Take Home	212/397	53.3%
Central Office Mail	116/244	47.4%

By May 10th, the cut-off date for data entry of questionnaires, a total of 775 questionnaires were processed. Considering all parents of children eligible to participate in the Student Survey, 363 of 440 (82.5 percent) of Student Take Home parents had completed parent questionnaires, while 279 of 402 (69.4 percent) of Central Office Mail parents had participated in the Parent Survey. Completion rates were then figured using a second formula, one which considered a base of only parents of children who had actually participated in Survey Day. This calculation was of interest since parents of children who had not participated in the Student Survey were not subject to Central Office and field follow-up on unreturned questionnaires. Here 365 of 415 (88 percent) of Student Take Home questionnaires were in, while 274 of 363 (75.5 percent) of Central Office Mail questionnaires had been received.

FINAL PARENT PARTICIPATION RATES WITH STUDENT TAKE HOME AND CENTRAL OFFICE MAIL (ELIGIBLE CHILDREN PARENTS AND SURVEYED CHILDREN PARENTS)

	ELIGIBLE		SURVEYED	
	number	percentage	number	percentage
Student Take Home	363/440	82.5%	365/415	88%
Central Office Mail	279/402	69.4%	274/363	75.5%

Criterion 3: quality and quantity of information provided. By the third criterion, quality and quantity of information provided, Student Take Home again emerged as distinctly superior. The first 609 parent questionnaires that arrived in house were subject to retrieval on critical items. Telephone calls were made to parents if inadequate answers were provided to selected important items on the questionnaire.

Two of the most personal critical items addressed locator information on the parent respondent (name, address, and telephone number) and the amount of the family's annual income. These two questions were chosen as the quality/quantity indicator for three reasons:

- the questions represent two different types of sensitivity
- both questions had high retrieval rates
- it seemed plausible that respondents could vary in their willingness to respond to these questions according to the manner in which they received a parent questionnaire.

Of 296 Student Take Home cases considered, 18 (6.2 percent) initially did not provide Respondent Locator Information and 34 (11.5 percent) did not provide Family Income Information. Of 177 Central Office Mail cases considered, 20 (11.3 percent) initially did not provide Respondent Locator Information and 41 (23.2 percent) did not provide Family Income Information.

INITIALLY MISSING CRITICAL INFORMATION WITH STUDENT TAKE HOME AND CENTRAL OFFICE MAIL

	RESPONDENT LOCATOR		INCOME	
	number	percentage	number	percentage
Student Take Home	18/296	6.2%	34/296	11.5%
Central Office Mail	20/177	11.3%	41/177	23.2%

On the basis of these results, in the NELS:88 Base-Year study, Student Take Home is being exclusively used as the method for delivering parent questionnaires. To enhance what appears to be the favorable impact of having the questionnaires sent from the school, we are regularly sending the parent questionnaires to the school two weeks prior to the student survey date and asking the school officials to hand out the parent packets. The parents have the option either of having the child deliver the questionnaire back to the school or of mailing the questionnaire directly back to NORC. It is often the case that the School Coordinator volunteers to track the return of the questionnaires. The method appears to be working exceptionally well.

The Second Experiment (and Some Ramifications): Identity of the Parent Respondent (and of the Eighth Grader's Various Other Parents/Guardians)

There are several reasons why it may be more convenient to mail a questionnaire to a pre-specified respondent rather than to the unspecific: TO THE PARENT OF: Name of Child. Among the most important of these is the possibility that the child's last name differs from that of the parent; for example, the mother has remarried and taken the new husband's name. This type of problem may at least partially explain why pre-prompting response rates were weak for the Central Office Mail treatment as compared to the Student Take Home treatment. Nevertheless, we felt it important that each of the child's parents be acknowledged as potentially the parent most knowledgeable concerning the eighth grader's education. We were curious whether the division of labor within American households would be such that the task of filling out a questionnaire dealing with the child's education would always be delegated to the mother. As will be illustrated below, we did achieve variation in gender of the parent respondents. As will be described in detail below, this initial inquiry led to a deeper investigation of several knotty problems related to the identity of the NELS:88 parents.

Who Filled Out the Parent Questionnaire?

Table 1 presents the distribution of "parent" respondents by relationship to the student both for parents of eighth graders in the 1987 NELS:88 Field Test and for parents of tenth and twelfth graders in the 1980 High School and Beyond Base Year Study. At the bottom of the table the manner of contacting the parents is described for each study.

TABLE 1. WHO FILLED OUT THE PARENT QUESTIONNAIRE?

	NELS:88 FIELD TEST	HS&B BASE YEAR	
		SENIOR	SOPHOMORE
Mother	583/775 (75%)	58.4%	60.8%
Stepmother	3/775 (< 1%)	1%	1%
Father	145/775 (18.7%)	36%	34.4%
Stepfather	6/775 (<1%)	1%	1.1%
Grandmother	8/775 (1%)	1%	<1%
Grandfather	3/775 (<1%)	<1%	<1%
Other Male Relative/ Guardian	2/775 (<1%)	2.3%	1.6%
Other Female Relative/ Guardian	7/775 (1%)		
Multiple Response	18/775 (2.3%)		

The NELS:88 Parent Questionnaire was addressed: To the Parents of: NAME OF CHILD. On the first page of the questionnaire, the parents were instructed: "The questionnaire should be completed by the parent or guardian who is most familiar with the student's current school situation and educational plans."

The HS&B Parent Questionnaire was addressed using the following technique: "From the student identification pages, we were able to obtain the parents' names and addresses. If the student reported both a male and female parent, a letter was sent to the female. If only one parent was mentioned, the letter was addressed to that parent."

It is interesting that, even though the HS&B questionnaire was usually mailed specifically to the mother, the percentage of mothers answering the questionnaire was much lower than it was in the NELS:88 Field Test. This is quite possibly explained by the fact that a good many financial questions appeared on the HS&B questionnaires, and some of these were quite complex. On the other hand, the NELS:88 questionnaire had only a few quite simple financial questions. It is likely that often the women just handed the questionnaire over to the husbands, who then filled out all of the questionnaire.

This set of figures is highly suggestive for at least two reasons: First, since the gender mix for NELS:88 Field Test and HS&B is so different, there may be a sex bias when comparing base-year results of the two studies. Second, if in 1992 there is a Parent Survey in the Second Follow Up for NELS:88 (as is anticipated in the research design), and if, too, a large number of complicated financial questions are in the questionnaire, then there might occur a large shift in respondent identity. That is, many of the mothers who filled out questionnaires in 1988 might have their spouse/partners fill out the 1992 questionnaires. If males tend to answer certain attitudinal or behavioral or perceptual items differently from females, this could be misinterpreted as a shift in attitude or behavior or perception of the child's parent individually and of the NELS:88 parents collectively.

As for the specific question of whether it is always the mother who will fill out a questionnaire concerned with the child's education, the answer is clearly negative. Focussing on the NELS:88 results, we see that only 75 percent of the time was it the mother who was the respondent. Nearly 19 percent of the time, the father filled out the questionnaire.

To gain some indication of whether our fear of sex bias (the tendency of male parents to systematically answer questions differently from female respondents) might be justified, we ran a number of crosstabulations on mother/step-mother and father/step-father respondents and their replies to various behavioral and attitudinal questions.

The following were found to be statistically significant at the indicated levels:

1. Do you speak a language other than English? (Significance: .0033)

2. Do you enforce rules for your child about any of the following activities?
   a. The grades s/he must maintain in school? (Significance: .0140)
   b. Doing homework? (Significance: .0135)
   c. Doing household chores? (Significance: .1005)

3. How many times during the past year has your child's school contacted you about doing volunteer work? (Significance: .1386)

Crosstabulation : Q1.22 DO YOU SPEAK A LANGUAGE OTHER THAN ENGLISH?
BY
Q1.1A R'S RELATIONSHIP TO STUDENT

Q1.1A-> / Q1.22	Count / Row Pct / Col Pct	MOTHER/ STEP-MOTHER 1	FATHER/ STEP-FATHER 2	Row Total
YES	1	220	77	297
		74.1%	25.9%	40.5%
		37.7%	51.0%	
NO	2	361	72	433
		83.4%	16.6%	59.0%
		61.9%	47.7%	
MISSING	9	2	2	4
		50.0%	50.0%	.5%
		.3%	1.3%	
	Column Total	583	151	734
		79.4%	20.6%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
11.45172	2	.0033	.823	2 OF 6 (33.3%)

Crosstabulation: Q4.5A DO YOU ENFORCE RULES FOR YOUR CHILD ABOUT THE GRADES HE/SHE MUST MAINTAIN IN SCHOOL?
BY
Q1.1A R'S RELATIONSHIP TO STUDENT

Q1.1A-> / Q4.5A	Count / Row Pct / Col Pct	MOTHER/ STEP-MOTHER 1	FATHER/ STEP-FATHER 2	Row Total
YES	1	455	105	560
		81.3%	18.8%	78.7%
		80.7%	70.9%	
NO	2	109	43	152
		71.7%	28.3%	21.3%
		19.3%	29.1%	
	Column Total	564	148	712
		79.2%	20.8%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
6.04058	1	.0140	31.596	None
6.60723	1	.0102	(Before Yates Correction)	

Crosstabulation: Q4.5B DO YOU ENFORCE RULES FOR YOUR CHILD ABOUT DOING HOMEWORK?
BY
Q1.1A R'S RELATIONSHIP TO STUDENT

Q4.5B / Q1.1A->	Count / Row Pct / Col Pct	MOTHER/ STEP-MOTHER 1	FATHER/ STEP-FATHER 2	Row Total
1 YES		496	124	620
		80.0%	20.0%	89.7%
		91.2%	84.4%	
2 NO		48	22	70
		68.6%	31.4%	10.1%
		8.8%	15.0%	
9 MISSING		0	1	1
		.0%	100.0%	.1%
		.0%	.7%	
Column Total		544	147	691
		78.7%	21.3%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
8.61133	2	.0135	.213	2 OF 6 (33.3%)

Crosstabulation: Q4.5C DO YOU ENFORCE RULES FOR YOUR CHILD ABOUT DOING HOUSEHOLD CHORES?
BY
Q1.1A R'S RELATIONSHIP TO STUDENT

Q4.5C / Q1.1A->	Row Pct / Col Pct	MOTHER/ STEP-MOTHER 1	FATHER/ STEP-FATHER 2	Row Total
1 YES		450	114	564
		79.8%	20.2%	82.5%
		83.8%	77.6%	
2 NO		87	33	120
		72.5%	27.5%	17.5%
		16.2%	22.4%	
Column Total		537	147	684
		78.5%	21.5%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
2.69730	1	.1005	25.789	None
3.11423	1	.0776	(Before Yates Correction)	

Crosstabulation: Q5.1E HOW MANY TIMES DURING THE PAST YEAR HAS YOUR CHILD'S SCHOOL CONTACTED YOU ABOUT DOING VOLUNTEER WORK?
BY
Q1.1A R'S RELATIONSHIP TO STUDENT

Q1.1A-> / Count / Row Pct / Col Pct	MOTHER/ STEP-MOTHER 1	FATHER/ STEP-FATHER 2	Row Total
Q5.1E			
0 NONE	336 77.1% 59.2%	100 22.9% 67.6%	436 60.9%
1 ONCE OR TWICE	150 82.4% 26.4%	32 17.6% 21.6%	182 25.4%
2 THREE OR FOUR	34 75.6% 6.0%	11 24.4% 7.4%	45 6.3%
3 MORE THAN FOUR	47 90.4% 8.3%	5 9.6% 3.4%	52 7.3%
9 MISSING	1 100.0% .2%	0 .0% .0%	1 .1%
Column Total	568 79.3%	148 20.7%	716 100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
6.94994	4	.1386	.207	2 OF 10 (20.0%)

While it is necessary to caution that there could be a self selection bias in operation -- perhaps fathers/stepfathers who chose to answer the questionnaire had special characteristics -- it should be noted that there were not significant differences by sex on the following questions:

1. How far in school do you expect your child to go? (Significance: .4907)

2. How satisfied are you with the education your child has received up to now? (Significance: .5655)

3. How many times during the past year has your child's school contacted you about each of the following?

a. Your child's academic work or choice of courses? (Significance: .9936)
b. Your child's behavior? (Significance: .8752)
c. Doing something for the school (like fundraising)? (Significance: .5964)

4. Indicate the extent of your agreement or disagreement with each of the following statements about the school your child attends:
a. My child is challenged at school. (Significance: .8415)
b. The school is preparing students well for high school. (Significance: .4909)

There seems to be some evidence here to suggest that a child's parents may not necessarily be collective in their opinions and perceptions, and that therefore, it does matter whether it is the male parent or female parent who fills out the questionnaire. The issue deserves further investigation in an experiment that has both parents of each child being asked all of the questions of the Parent Questionnaire.

Returning briefly to the matter of the male parent being more likely to fill out the Parent Questionnaire if the questionnaire contains a substantial number of financial items, it is interesting to note that the post-retrieval non-response rate on the Field Test item concerning "total family income received in the past year" was 11.2 percent for mothers/stepmothers and 5.4 percent for fathers/stepfathers. A high non-response rate for females on the item may indicate lack of knowledge about finances. Even if it should turn out that mothers do usually know more about the child's education, division of labor within the household may mean that the father is the expert on family finances. If it is important to have accurate information about the family's financial situation -- and it is -- then the implications of gathering such information from a second-best respondent should be carefully examined. To complicate matters further, over time the traditional household roles may be breaking down. It is possible that the mothers of 1992 will be more savy about family finances than were their counterparts in 1980. Obviously this has potential importance for longitudinal research.

Amount of Time Child Lives With the Parent

One concern we had about not sending the Parent Questionnaire to a pre-designated parent was that the child who lived part of the time with someone other than a custodial parent might deliver the questionnaire to the non-custodial parent. While we did want the questionnaire to be answered by the parent who was most familiar with the child's educational experiences, for purposes of studying the socio-economic status of the child, we also hoped to have as the respondent a parent with whom the child lived most of the time. In asking the school personnel for addresses, we specifically asked for the mailing address of the parent with whom the child was living. When handing out the questionnaires to the children to carry home, we told the children that if their parents lived in two different places, they should give the questionnaire to the parent who lived in the place that

they considered to be home. Table 2 presents results that allay our fears. Out of 775 cases, only one parent each reported the child living in the home less than half the time or none of the time.

TABLE 2. AMOUNT OF TIME CHILD LIVES WITH THE RESPONDENT:

All of the time	761/775 (98.2%)
> 1/2 time	11/775 (1.4%)
< 1/2 time	1/775 (<1%)
None of the time	1/775 (<1%)
Blank	1/775 (<1%)

Who the Child Reports Living With

The first column of Table 3 (see next page) shows which adults live in the same households with the NELS:88 eighth graders. Of special interest is the fact that nearly 33 percent of the eighth graders do not reside with both of their biological parents. A full 16 percent reside with only their mother, more than 8 percent live with their mother and a male guardian, nearly another 3 percent live with the mother and grandparents. Only about 2.6 percent of the eighth graders reside with their biological father, with or without a female guardian.

Since the NELS:88 Field Test sample was not a full national probability sample, in the second and third columns of Table 3, we present comparable statistics obtained from a 1981 study conducted for the National Center for Health Statistics to demonstrate that in fact the percentages found in the NELS:88 sample do not deviate greatly from national proportions for this age group.

Two considerations are particularly important here. First, in at least 33 percent of the cases, there is a potential source of confusion between the parent respondent's spouse/partner and the child's other biological parent. For the NELS:88 Base Year Parent Questionnaire, in order to clarify such reference, the parent respondent has been given strict instructions concerning who should be considered the spouse/partner (the person living in the household with the parent respondent). Because it is more difficult both to describe to an eighth grader what is needed and to successfully elicit from an eighth grader information on the mother's or father's spouse/partner as opposed to the child's natural parent, one cannot have complete confidence that the child will describe, when filling out his/her questionnaire, the same individual that his/her parent describes when filling out the Parent Questionnaire. What this means is that the mother might say that her spouse/partner is a banker, while the child might say that his father is a farmer. Both are reporting accurately,

TABLE 3. WHO THE CHILD REPORTS LIVING WITH:

In the Field Test Student Questionnaire students were asked: WHICH OF THE FOLLOWING PEOPLE LIVE IN THE SAME HOUSEHOLD WITH YOU? A list of various relationships and nonrelationships followed. Examination of a random sample of 857 of the student questionnaires resulted in the following summary:

	NELS:88	NCHS (ALL)	NCHS (12-17)
Mother/Father	576/857 (67.2%)	[67.5%]	61.9%
Mother	137/857 (16%)	[18.3%]*	18.8%
Mother/Male Guardian	71/857 (8.3%)	[7.1%]	9.7%
Mother/Grandparents	25/857 (2.9%)	*	
Father:	8/857 (.9%)	[1.5%]	2.3%
Father/Female Guardian	16/857 (1.86%)	[1.6%]	2.3%
Other:	24/857 (2.8%)		
Father/ Other Male Guardian	1		
Other Male Guardian/ Other Female Guardian	5		
Father/ Grandparents	3		
Mother/ Other Female Guardian	1		
Brother(s)/ Sister(s)	2		
Grandparents/Other Relatives	4		
Grandparents	6		
Other Female Guardian	2		

The NELS:88 Field Test sample is not a full national probability sample; the field test was conducted in only five states; private schools were overrepresented as were Hispanics and Asian-Pacific Islanders. It is interesting to note, therefore, how close these percentages are to those cited by Nicholas Zill in his 1987 study of stepfamilies. His data came from a 1981 study conducted by the National Center for Health Statistics which surveyed 15,416 children aged 0-17. His figures appear in square parentheses to the right of the NELS:88 figures. To the right of those are the figures for the 12-17 year old subsample.

* When the Mother/Grandparents figures are added in, the equivalent NELS:88 figure is 18.9%.

but certainly comparative SES measures could appear confusing by these results if the investigator was not warned of the source of the disparity.

To examine the extent to which we should be concerned about such potential confusion, we conducted two sets of crosstabulations using data from both the Student and Parent Surveys.

The first examination was done as follows:

1. In the Parent Questionnaire, the question was asked: "Is there one parent or are there two parents in your household?" We separated out those who answered "one parent".

2. Also in the Parent Questionnaire, the question was asked: "What is your relationship to the student?" We distinguished those responding "mother" from those answering "father".

3. Then, using data from the Student Survey, we ran the frequencies for answers the children of these single parents provided on three questions concerning the parent of the opposite sex:

   a. Which of the categories below comes closest to describing the job of your father (or male guardian) — if single parent was the mother, or mother (or female guardian) -- if single parent was the father.

   RESULTS:

   When the mother was the single parent, 80 of 161 children (49.7%) provided an answer to the question; another 25 children (15.5%) gave multiple answers (marked more than one occupation code when only a single code was considered valid); 26 (16.1%) answered "Don't Know"; 29 (18%) did not answer at all, and 1(.6%) refused to answer. When the father was the single parent, 13 of 17 (76.5%) children gave a single answer concerning the mother, while 3 (17.6%) answered "Don't Know" and 1 (5.9%) did not answer at all.

   b. What was the highest level of education your father (or male guardian) and your mother (or female guardian) completed?

   RESULTS:

   When the mother was the single parent, 76 (47.2%) of the children gave single answers, 47 (29.2%) gave (invalid) multiple response answers, and 38 (23.6%) answered "Don't Know". When the father was the single parent, 8 children (47.1%) gave one answer concerning the mother, while 4 (23.5%) answered "Don't Know," and 5 (29.4%) gave an invalid multiple response.

These results should be considered with care. At a minimum, they suggest that we are receiving a great deal of information that does not pertain to the immediate household. But the question that is immediately suggested is the following: Do we really want to tell the

child: Don't answer about your other birth parent if that parent doesn't live in your house with you? In considering socio-economic measures in a study purporting to examine the influences on a child's educational development, surely in the ideal case, we would want to know about the total socio-economic impact on, say, the formation of the child's academic aspirations. The child and the respondent parent may differ in what they can and should contribute to letting us understand the full set of socio-economic influences. If the child's mother is living in a low socio-economic status following divorce from the child's better financially situated biological father, surely we want to know about her income, not about that of her ex-spouse. On the other hand, if a child's perception of the biological father's expectations with regard to the child's educational attainment importantly determine such attainment, neither do we want to fail to register this influence.

Consider the results of the next question examined:

c. How far in school do you think your parents (guardians) want you to go? Circle one answer for father (other male guardian), another answer for mother (female guardian).

RESULTS: When the mother was the single parent, only 18 of the 161 students answered the question concerning the absent parent with "Don't Know" (17 or 10.6%) or left the item blank (1 or .6%). The others answered over a range of grade levels from Graduate From High School to Attend a Higher Level of School After Graduating From College. When the father was the single parent, only one of the 17 students responded "Don't Know". The others provided answers.

Thus it would seem necessary to explore the question: If the students have a perception of an education level that their parent living outside of the house wants them to achieve, what is the extent of the influence these parents have on their children's educational aspirations and achievements?

Still pursuing the matter of possible confounding of parental identity, we went on to examine a second set of crosstabulations as follows:

1. In the Parent Questionnaire, the question was posed: "What is your current marital status?" We separated out the cases in which the response was "married."

2. In the Locator Booklet of the Student Questionnaire, students were asked: "Which of the following people live in the same house with you?" A list of various relationships followed. For the students whose parents were married, we pulled the cases in which the child had listed "Other male guardian (stepfather or foster father)" and "Other female guardian (stepmother or foster mother)."

3. Also in the Locator Booklet of the Student Questionnaire, students were first asked to list the name of the mother or female guardian, then were asked to check a box if the address and

telephone number of the mother or female guardian was "Same as Mine". If the student did not check the box, s/he was asked then to give the address and telephone of the mother or female guardian. A comparable question was asked concerning the father or male guardian. We were interested in seeing the extent to which the child did not check the YES box. That is, we hoped to have a rough estimate as to the extent that the child gave information for the birth parent living outside of the house instead of for the step-parent living in the house.

RESULTS: When the child lived with a stepfather (and presumably the birth mother), in 15 of 52 such cases (28.8 percent of the time), the child did not check the YES box. Since only 14 children lived with their father and a stepmother, we chose not to consider this case. However, to have some grasp of the extent that the children might just be forgetting to mark the box, we looked at the responses the 52 mother-stepfather cases gave to the mother/female guardian box. Here 6 responses were not checked. This would lead us to believe that 9 of the cases described above could have been instances in which the child was answering with respect to the out-of-house birth father.

As important as keeping track of which parent (biological or step) the eighth grader will choose as the questionnaire referent is determining which of the two parent types may be most influential for a given child. A parent who has only been in the household for a year or so could potentially have little influence on outcomes of the child's high school years and thereafter. Similarly, a parent who has just recently departed the household could have substantial influence. Or both types of parents could have separate and unique influences on the student. If one assumes that parents have a profound influence on the child's development, and this is generally conceded to be a reasonable assumption, then it would seem wise to get a grasp of the child's perceptual identity of "his/her" "parents" and then to measure the full range of contributions of those parents to the child's educational development.

Current Marital Status of Respondent

To this point, we have suggested that comparing NELS:88 results to HS&B results or comparing NELS:88 results to Second Follow-Up results (1992) or even comparing NELS:88 Student Survey results with NELS:88 Parent Survey results could be confounded by an overall slipperiness in identification of the "parents". Table 4, which presents the current marital status of the NELS:88 Parent Survey respondents, further substantiates the validity of this concern. By the Field Test results, a full 24 percent of the eighth graders are living in a household that does not have both a male and a female parent.

Also presented in Table 4 are comparable results for the HS&B (1980) parents (those of both senior and sophomore cohorts) and those of the 1976 National Survey of Children. The latter two studies were full national probability studies. Controlling for the fact that the NSC results are for parents whose children are between 7 and 11 (thus there has been less time for divorce to occur than for parents of 13

TABLE 4. CURRENT MARITAL STATUS OF RESPONDENT:

	NELS:88	HS&B (1980) Senior	Sophomore	1976 National Survey of Children
Single, never Married	30/765 = 3.9%	1.1%	1.1%	2.28%
Married	581/765 = 75.9%	79.6%	78.3%	82.24%
Separated	35/765 = 4.6%	4.1%	4.5%	5.6%
Divorced	95/765 = 12.4%	9.3%	10.5%	7.8%
Widowed	24/765 = 3.1%	5.2%	4.6%	2.02%

The 1976 National Survey of Children was a multi-stage stratified probability sample of households in the continental United States containing at least one child born between 9/1/1964 and 12/31/1969 (a child between the ages of 7 and 11). Since most eighth-graders are 13 years old, the NSC is obviously not precisely comparable. Nevertheless, the above figures suggest a substantial rise in the divorce rate and a comparable drop in the marriage incidence.

Since the NELS:88 Field Test Parent Survey, the HS&B Study, and the 1976 National Survey of Children all oversampled particular ethnic populations and since the ethnic groups have been shown to vary greatly in their incidences of the various marital status categories, below we provide the ethnic breakdowns of the parent respondents for the three studies.

NELS:88 FIELD TEST	1976 National Survey of Children	HS&B Senior	Sophomore
Black 15.5%	Black 14.5%	Black 12.3%	14%
White 69.5%	White 75.5%	White 70%	70.6%
Asian/Pacific Islander 5.5%	Oriental 5.8 %	API 1.9%	1.5%
Other 7.1%	Other 4.5%	Hispanic 9.3%	10.4%
		Other 6.5%	3.5%

year olds), while the HS&B results are for parents of older students, it appears that over time there has been a substantial rise in the divorce rate and a comparable drop in the marriage incidence. In a four year time span, we might expect considerable change in the composition of the "parents" of the NELS:88 eighth graders.

Relevant questions are the following: When Mom reports information about her spouse/partner in 1992, will she be referring to the same individual who occupied that role in 1988? To what extent does the fact that a divorce or a remarriage occurs between the time a child is an eighth grader and the time s/he is a senior in high school alter expectations for the child's educational achievement? Again, who is likely to have the greater influence on a child, the parent who helped to raise the child most of his/her life, or a person who enters the family as a step parent when the child is well into the teenage years? To what extent is it dangerous -- or justifiable -- to totally ignore the psychological, financial, other socio-economic influence of a non-custodial parent? These are all questions that might be fruitfully addressed in future follow-ups of NELS:88.

SUMMARY

In sum, data gathered from the NELS:88 Parent Survey Field Test suggest that more investigation is required on how best to measure parental influence on student educational achievement. Asking just one of the student's parents to provide information seems not to be sufficient, particularly when two parents occupy different roles in the household and thereby acquire different perceptions of the student's education and the family's situation. Moreover, in the not infrequent case in which a child has a parent living outside of the household, exclusive focus of research on the child's residential home may also be insufficient. Especially in longitudinal studies concerned with educational development, it would seem imperative that attention be paid to the reality of multiple parents and multiple parent opinions. At very least, substantial care needs to be taken to pin down the identity of all relevant players. It is equally important to learn the means by which a child's parents influence educational development over time.

REFERENCES CITED

Ingels, Steven et al., Field Test Report: National Education Longitudinal Study of 1988 (NELS:88), 1987.

National Opinion Research Center, High School and Beyond: Parent Questionnaire Codebook, Report to National Center for Education Statistics, August 1981.

National Survey of Children (1976): unpublished statistical frequencies of the study conducted by the Institute for Survey Research at Temple University.

Zill, Nicholas, "Behavior, Achievement, and Health Problems Among Children in Stepfamilies: Findings From A National Survey of Child Health," 1987. Scheduled to appear in E. Mavis Hetherington and I. Arasten (Eds.), The Impact of Divorce, Single Parenting, and Step-parenting on Children.